RESEARCH REPORT

HEALTH EFFECTS INSTITUTE

Number 94, Part I
June 2000

Final Version

The National Morbidity, Mortality, and Air Pollution Study
Part I: Methods and Methodologic Issues

Jonathan M Samet, Francesca Dominici, Scott L Zeger,
Joel Schwartz, Douglas W Dockery

Includes a Commentary by the Institute's Health Review Committee

HEALTH
EFFECTS
INSTITUTE

The Health Effects Institute, established in 1980, is an independent and unbiased source of information on the health effects of motor vehicle emissions. HEI sponsors research on all major pollutants, including regulated pollutants (such as carbon monoxide, ozone, nitrogen dioxide, and particulate matter) and unregulated pollutants (such as diesel engine exhaust, methanol, and aldehydes). To date, HEI has supported more than 200 projects at institutions in North America and Europe and has published over 100 research reports.

Typically, HEI receives half its funds from the US Environmental Protection Agency and half from 28 manufacturers and marketers of motor vehicles and engines in the US. Occasionally, funds from other public and private organizations either support special projects or provide resources for a portion of an HEI study. Regardless of funding sources, HEI exercises complete autonomy in setting its research priorities and in reaching its conclusions. An independent Board of Directors governs HEI. The Institute's Research and Review Committees serve complementary scientific purposes and draw distinguished scientists as members. The results of HEI-funded studies are made available as Research Reports, which contain both the Investigators' Report and the Review Committee's evaluation of the work's scientific quality and regulatory relevance.

HEI STATEMENT
Synopsis of Research Report 94, Part I

The National Morbidity, Mortality, and Air Pollution Study: Methods and Methodologic Issues

BACKGROUND

Epidemiologic time-series studies conducted in a number of cities have found, in general, an association between daily changes in particulate matter (PM) and daily mortality counts. These studies, which have raised concerns about public health effects of particulate air pollution, have contributed significantly to decisions about regulating PM in the US. However, scientists have pointed out a number of limitations in these studies that raise questions about the interpretation of the results. They have questioned whether PM actually causes increased mortality, pointing to differences in results among studies in different locations; inadequate consideration of other variables, such as other pollutants, that might account for such an association; and the use of different analytic methods, yielding different results. Others have also pointed to the fact that the time-series epidemiologic studies do not include measurements of each individual's exposure to air pollutants, and concentration data from area monitors in the city are used as an approximation instead. This use of a surrogate may introduce error in exposure measurement that could account for observed associations.

Scientists have also suggested that the PM-mortality association represents premature mortality by only a few days among those near death. Advancing of death by only a few days has been referred to as *harvesting* or *mortality displacement*. If associations between increased mortality and PM reflect solely short-term mortality displacement, the daily time-series studies may be showing an effect of limited public health impact.

As the US Environmental Protection Agency (EPA) prepares to reevaluate the ambient air standards for PM in the US, and as other countries are similarly reviewing the evidence for associations between PM and mortality, it is important to understand whether any observed associations might be accounted for by other pollutants, whether bias in exposure measurements leads to the association, or whether mortality displacement underlies the association.

APPROACH

In an effort to address the uncertainties regarding the association between PM and daily mortality, and to determine the effects of other pollutants on this association, HEI funded the National Morbidity, Mortality, and Air Pollution Study (NMMAPS). Dr Jonathan Samet and his colleagues at Johns Hopkins University, in collaboration with investigators at Harvard University, conducted this time-series study in large cities across the US where levels of PM and gaseous pollutants were varied. To conduct such a study and to begin to address issues of exposure measurement error and mortality displacement, new analytic methods were needed. This first report includes 5 separate sections that describe such methods; a second report will describe the results from applying these methods.

RESULTS AND IMPLICATIONS

Exposure Measurement Error

Dr Samet and his colleagues have advanced our understanding of the effects of error in measuring pollution in time-series studies. They use a theoretical model to test systematically what effect the relationship between personal exposure and ambient exposure might have on the observed increase in mortality associated with PM. The application of this model for correction requires using both ambient monitoring data and some personal exposure measurement data in the same area. Such availability of both types of data is currently limited to a few locations. The investigators were able to apply one such set of personal and ambient exposure data as an example. The theoretical and actual analyses generally appear to refute the criticisms that exposure measurement error could explain the associations between PM and adverse health effects. The general absence of measured exposure data, however, precludes making firm conclusions as to the specific effects of the errors. HEI, EPA, and other organizations are currently funding research to obtain more exposure data that should lead to more confident specific conclusions regarding the effect of any exposure measurement error.

This Statement, prepared by the Health Effects Institute, summarizes a research project sponsored by HEI since 1996 and conducted by Dr Jonathan M Samet of Johns Hopkins University School of Public Health, Baltimore MD. The following Research Report contains the detailed Investigators' Report, Part I, and the Commentary on Part I prepared by a panel of the Institute's Health Review Committee.

Mortality Displacement

The investigators' examination of the role of mortality displacement using 2 different but related statistical approaches is original and suggests that more than a short-term displacement of mortality is occurring. It remains unclear, however, which component(s) of the air pollution mix are actually responsible for any longer-term effect and, given our inexperience with the methods that focus on the longer time scales, how the estimates of effect based on these methods should be interpreted. Broader application of these methods and the development of new methods are needed to understand the public health implications more clearly.

Multicity Analysis Methods

The analytic methods developed to examine multicity mortality are flexible and comprehensive, allowing a combination of mortality effects across cities and an exploration of factors that might account for differences in effects seen among individual cities. These methods also set the stage for further analyses and for addressing questions that might follow from initial results.

Library of Congress Catalog Number for the HEI Report Series: WA 754 R432.
The paper in this publication meets the minimum standard requirements of the ANSI Standard Z39.48-1984 (Permanence of Paper) effective with Report Number 21, December 1988, and with Report Numbers 25, 26, 32, 51, 65 Parts IV, VIII, and IX, and 91 excepted. These excepted Reports are printed on acid-free coated paper.

HEALTH EFFECTS INSTITUTE

CONTENTS

Research Report 94

The National Morbidity, Mortality, and Air Pollution Study
Part I: Methods and Methodologic Issues

HEI STATEMENT

This Statement, prepared by the Health Effects Institute, is a nontechnical summary of the Investigators' Report and the Health Review Committee's Commentary.

PREFACE

INVESTIGATORS' REPORT

When an HEI-funded study is completed, the investigators submit a final report. The Investigators' Report is first examined by three outside technical reviewers and a biostatistician. The Report and the reviewers' comments are then evaluated by members of the HEI Health Review Committee, who had no role in selecting or managing the project. During the review process, the investigators have an opportunity to exchange comments with the Review Committee and, if necessary, revise the report.

Overview of Study Design and Conclusions

Jonathan M Samet, Francesca Dominici, Scott L Zeger, Joel Schwartz, and Douglas W Dockery

Departments of Epidemiology and Biostatistics, Johns Hopkins School of Public Health, Baltimore, Maryland, and Department of Environmental Health, Harvard School of Public Health, Boston, Massachusetts

Section 1: Exposure Measurement Error in Time-Series Studies of Air Pollution

Scott L Zeger, Duncan Thomas, Francesca Dominici, Jonathan M Samet, Joel Schwartz, Douglas W Dockery, and Aaron Cohen

Departments of Biostatistics and Epidemiology, Johns Hopkins School of Public Health, Baltimore, Maryland; Department of Preventive Medicine, University of Southern California School of Medicine, Los Angeles, California; Department of Environmental Health, Harvard School of Public Health, Boston, Massachusetts; and Health Effects Institute, Cambridge, Massachusetts

continued

Section 2: A Measurement Error Model for Time-Series Studies of Air Pollution and Mortality

Francesca Dominici, Scott L Zeger, and Jonathan M Samet

Departments of Biostatistics and Epidemiology, Johns Hopkins School of Public Health, Baltimore, Maryland

Section 3: Mortality Displacement–Resistant Estimates of Air Pollution Effects on Mortality

Scott L Zeger, Francesca Dominici, and Jonathan M Samet

Departments of Biostatistics and Epidemiology, Johns Hopkins School of Public Health, Baltimore, Maryland

Section 4: Mortality Displacement and Long-term Exposure Effects Related to Air Pollution and Mortality

Joel Schwartz

Department of Environmental Health, Harvard School of Public Health, Boston, Massachusetts

Section 5: Combining Evidence on Air Pollution and Daily Mortality from Twenty Largest US Cities

Francesca Dominici, Jonathan M Samet, and Scott L Zeger

Departments of Biostatistics and Epidemiology, Johns Hopkins School of Public Health, Baltimore, Maryland

continued

COMMENTARY Health Review Committee

The Commentary about the Investigators' Report is prepared by the HEI Health Review Committee and staff. Its purpose is to place the study into a broader scientific context, to point out its strengths and limitations, and to discuss the remaining uncertainties and the implications of the findings for public health.

RELATED HEI PUBLICATIONS: PARTICULATE MATTER

Publishing History: This document was posted as a preprint on *www.healtheffects.org* on 5/15/00 and finalized for print on 9/7/00.

Citation for whole report:

Samet JM, Dominici F, Zeger SL, Schwartz J, Dockery DW. 2000. The National Morbidity, Mortality and Air Pollution Study, Part I: Methods and Methodologic Issues. Research Report 94. Health Effects Institute, Cambridge MA.

When citing a section of this report, refer to it as a chapter of this document.

PREFACE

ORIGINS AND OBJECTIVES

In 1996, HEI initiated the National Morbidity, Mortality, and Air Pollution Study (NMMAPS)*, based on the realization that a national study could address one of the major questions regarding air pollution and daily mortality: whether particulate air pollution is responsible for the associations between air pollution and daily mortality that have been observed in multiple studies, or whether these associations are due, in part or completely, to other air pollutants. This realization emerged both from the experience of the Particle Epidemiology Evaluation Project (PEEP), funded by HEI from 1994 to 1997, and from an evaluation of the literature at that point, which largely included studies of single cities.

PEEP was designed to (1) address the replicability and validity of key epidemiologic studies of particulate air pollution and daily mortality by conducting detailed reanalyses of selected data sets, and (2) explore in more extensive data sets some of the larger scientific and public health issues raised by the findings of these earlier epidemiologic studies. PEEP investigators, led by Drs Jonathan Samet and Scott Zeger, successfully replicated the numerical results of the earlier studies, including the previously reported associations between total suspended particles (TSP) and daily mortality in Philadelphia (Schwartz and Dockery 1992). More detailed analysis of the Philadelphia data led Samet and Zeger to conclude, however, that the associations with air pollution in that city could not be attributed to particulate air pollution alone. In its Commentary on the contributions and limitations of PEEP (Samet et al 1997), the Oversight Committee concluded: "Although individual air pollutants (TSP, SO_2, and ozone) are associated with increased daily mortality [in Philadelphia], the limitations of the … data make it impossible to establish that particulate air pollution alone is responsible for the widely observed associations between increased mortality and air pollution in that city. All we can conclude is that it appears to play a role. … Ultimately, it will require joint analyses of data from multiple cities with different copollutant correlations … to address further the role of multiple pollutants."

NMMAPS was also designed to address two additional issues that complicated interpretation of the results of daily mortality and air pollution studies considered in PEEP: the effect of measurement error in exposure estimates on relative risk estimates, and whether any effect of life-shortening (mortality displacement) associated with increased daily mortality can be removed from estimates of risk associated with air pollution. With regard to exposure measurement error, the Oversight Committee stated in its Commentary on PEEP that "Errors in exposure measurements as a result of using data provided by centrally located monitors rather than exposures or doses measured in individuals, could, in the context of complex multivariable models for daily mortality, affect the relative risk estimates in ways that are difficult to predict. The possibility of such errors are an important source of uncertainty about the true magnitude of the estimated effects of individual air pollutants on daily mortality." For this reason, the Oversight Committee recommended "… developing models to assess exposure measurement errors in daily time-series analyses, and applying those models to a national data set using more detailed exposure data, if available."

The extent of life-span reduction associated with pollution-related daily mortality in Philadelphia and other locales remained unclear. If such reductions were small, due mainly to the advancement of the date of death for frail individuals by a matter of days (mortality displacement), then the public health implications would be less profound. The Oversight Committee remarked that "Estimating the extent of life-shortening caused by short-term elevations in air pollution remains one of the most important tasks for future studies." Developing methods for addressing the questions of whether any excess daily mortality is associated with air pollution only, or of whether any association largely reflects short-term mortality displacement, became an important methodologic objective of NMMAPS.

To address these questions, NMMAPS had the following two broad objectives:

- To conduct a nationwide study of acute health effects of air pollution on morbidity and mortality. NMMAPS is based on data from the US national air monitoring network provided by the US Environmental Protection Agency's (EPA's) Aerometric Information Retrieval System (AIRS) database, which contains information on particulate matter less that 10 μg in aerodynamic diameter (PM_{10}) and other criteria pollutants from 1987 to 1994, as well as from information on health and the population from the National Centers for Health Statistics, the Health Care Financing Administration, and the US Census. NMMAPS evaluates two issues: (1) air pollution and daily mortality in the 20 and

* A list of abbreviations and other terms appears after the Investigators' Report.

90 largest US cities, and; (2) daily hospital admissions of the elderly (≥ 65 years old) in 14 US cities with daily measurements of PM_{10}. A combined analysis using daily mortality *and* hospital admissions in the same cities is planned.

- To develop the statistical and epidemiologic methods required for data analysis and interpretation of results from such an investigation. NMMAPS investigators have developed methods for combining the evidence across multiple locations and for assessing the impact of exposure misclassification on the estimated association between daily mortality and air pollution. They have also developed approaches that begin to answer the question of whether or not the excess daily mortality that has been associated with air pollution reflects only, or largely, small reductions in survival among frail individuals.

NMMAPS focuses on the acute health effects of particulate air pollution, measured as PM_{10}. Its design, however, was intended by the investigators also to provide a framework for the study of pollutants other than particles.

PARTICIPANTS AND CONDUCT

NMMAPS has been conducted by a team of investigators from the Johns Hopkins School of Public Health, led by Principal Investigator Jonathan Samet and including Drs Scott Zeger and Francesca Dominici. As discussed above, Samet and Zeger had conducted PEEP, from which NMMAPS developed. The Johns Hopkins investigators were responsible for the design and analysis of the mortality component of NMMAPS. They have worked in collaboration with Drs Douglas Dockery and Joel Schwartz of the Harvard School of Public Health on methods for addressing mortality displacement and measurement error. Dockery and Schwartz designed and conducted the morbidity analyses.

NMMAPS has been overseen by the same Oversight Committee that worked on PEEP, on HEI's behalf. This committee, chaired by Dr Gerald van Belle of the University of Washington, comprises leading experts in epidemiology, biostatistics, pulmonary medicine, and aerometric measurement. The Oversight Committee was responsible for working with the investigators to develop, and ultimately to approve the analytic plan that has guided NMMAPS from its inception.

As the analytic plan for NMMAPS was being developed, HEI sought the comments of a broad range of scientists and technical experts from industry, government, and public interest groups. To provide continuing updates on the progress of the study to these diverse groups, HEI has organized regular presentations of interim results at its Annual Conference (1997 to 1999), a symposium at the International Society for Environmental Epidemiology (September 1999), and briefings for HEI sponsors (July 1997, February 1998, and December 1998). Besides providing interested parties with up-to-date information on the progress of NMMAPS, these events provided HEI, the Oversight Committee, and the investigators with valuable comments and suggestions for their work.

REPORT REVIEW

All HEI reports are reviewed by the HEI Health Review Committee and external reviewers with relevant expertise as required by the subject matter of the report. NMMAPS was reviewed by a Panel that included members of the HEI Health Review Committee as well as several other individuals with expertise relevant to the methods and analyses in this report. The Panel also wrote the Commentaries for Part I and Part II of the NMMAPS report with input from the full HEI Health Review Committee, members of the NMMAPS Oversight Committee, and the HEI Research Committee.

STRUCTURE OF THE HEI REPORT

The results of NMMAPS are presented as two reports. *Part I: Methods and Methodologic Issues* comprises a collection of methodologic papers on three topics: (1) measurement error in air pollution exposure, (2) mortality displacement, and (3) methods for combining the evidence in multiple locations using Bayesian hierarchical models. *Part II: Morbidity, Mortality, and Air Pollution in the United States* presents the results of analyses of daily mortality in the 20 and 90 largest US cities and in hospital admissions of the elderly (those 65 years old or older) in 14 US cities.

ACKNOWLEDGMENTS

HEI thanks the investigators and many other individuals whose contributions enhanced the quality of the NMMAPS project and this Research Report. Oversight of this complex project and evaluation of the findings would not have been possible without the help of the NMMAPS Oversight Committee, which included members of the HEI Research Committee, and the NMMAPS Review Panel, which included members of the HEI Review Committee. In

particular, the Institute thanks Dr Aaron Cohen for his role in assisting the Oversight Committee with management of this project, and Dr Diane Mundt for her role in assisting the Review Panel in developing the Commentary. The Review Panel gratefully acknowledges the cooperation of the investigators during the review process. Finally, the Institute appreciates the efforts of HEI's editorial and publications staff in preparing this Research Report.

INVESTIGATORS

Johns Hopkins School of Public Health

Jonathan M Samet (Principal Investigator)

Francesca Dominici

Scott L Zeger

Harvard School of Public Health

Douglas W Dockery

Joel Schwartz

REFERENCES

Samet JM, Zeger DL, Kelsall JE, Xu J. 1997. Air pollution, weather, and mortality in Philadelphia 1973–1988. In: Particulate Air Pollution and Daily Mortality: Analyses of the Effects of Weather and Multiple Air Pollutants, The Phase I.B Report of the Particle Epidemiology Evaluation Project. Health Effects Institute, Cambridge MA.

Schwartz J, Dockery DW. 1992. Increased mortality in Philadelphia associated with daily air pollution concentrations. Am Rev Respir Dis 145:600–604.

INVESTIGATORS' REPORT

The National Morbidity, Mortality, and Air Pollution Study
Part I: Methods and Methodologic Issues

Jonathan M Samet, Francesca Dominici, Scott L Zeger, Joel Schwartz, and Douglas W Dockery

Overview of Study Design and Conclusions

THE NMMAPS PROJECT

The National Morbidity, Mortality, and Air Pollution Study (NMMAPS)* comprises a comprehensive set of analyses of air pollution, mortality, and morbidity in a national sampling frame based on the monitoring information maintained in the US Environmental Protection Agency's (EPA's) Aerometric Information Retrieval System (AIRS). The project is a collaboration between investigators at Johns Hopkins School of Public Health (Drs Samet, Zeger, and Dominici) and Harvard School of Public Health (Drs Dockery and Schwartz). The project's overall objectives lie in the complementary domains of methods development and methods application. This report, the first of two parts, details the methodologic components of NMMAPS. Part II provides the substantive findings on air pollution, mortality, and morbidity (Samet et al 2000).

The objectives for developing specific methodologic components for NMMAPS are fivefold.

1. To develop semiautomated or automated approaches for database construction using databases of the EPA, the National Center for Health Statistics (NCHS), the Health Care Financing Administration (HCFA), the Census Bureau, and the National Weather Service;
2. To develop and apply statistical methods for regression analyses of the multisite data, and to develop spatial time-series methods to estimate spatial maps of the relative rates of mortality associated with air pollution, while accounting, as necessary, for the spatial and temporal correlations in the mortality data;
3. To develop and apply methods that adjust for smooth trends and seasonality on mortality caused by changing demographics and health behaviors, influenza epidemics, and other unidentified factors;
4. To examine the consequences of measurement error in the exposure variables for assessing pollutant-mortality associations; and
5. To examine the degree to which pollution-related mortality reduces years of life (mortality displacement).

The objectives for application of methods developed for NMMAPS are threefold.

1. To assess the relation between air pollution and mortality in the largest US cities monitored for PM_{10} from 1987 forward;
2. To assess the relation between air pollution and morbidity in selected US cities monitored for PM_{10} from 1987 forward; and
3. To conduct paired analyses of morbidity and mortality in the same locations.

The design for NMMAPS builds on prior work supported by the Health Effects Institute in the Particle Epidemiology Evaluation Project (PEEP) (Samet et al 1995, 1997). This project was initiated in 1994 with the objectives of validating the data and replicating the findings in several of the time-series studies of air pollution and mortality reported during the 1990s. In a second phase, PEEP addressed several methodologic issues. These included selecting the

* A list of abbreviations and other terms appears on page 13.

The National Morbidity, Mortality and Air Pollution Study: Methods and Methodologic Issues, Part I of Health Effects Institute Research Report 94, includes a Preface, an Investigators' Report, a Commentary by the Health Review Committee and an HEI Statement about the research project. Correspondence concerning the Investigators' Report may be addressed to Dr Jonathan M Samet, Department of Epidemiology, Johns Hopkins School of Public Health, 615 North Wolfe Street, Ste W 6041, Baltimore MD 21205-2179.

Although this document was produced with partial funding by the United States Environmental Protection Agency under Assistance Award R824835 to the Health Effects Institute, it has not been subjected to the Agency's peer and administrative review and therefore may not necessarily reflect the views of the Agency, and no official endorsement by it should be inferred. The contents of this document also have not been reviewed by private party institutions, including those that support the Health Effects Institute; therefore, it may not reflect the views or policies of these parties, and no endorsement by them should be inferred.

approach for controlling for potential confounding by weather (Samet et al 1998) and determining the sensitivity of findings to model-building strategies (Kelsall et al 1997; Samet et al 1997).

The present project, NMMAPS, evolved from PEEP. The objectives encompassed methodologic issues that were persistent sources of uncertainty in interpreting the epidemiologic evidence: mortality displacement and exposure measurement error. The plan for multicity analyses was prompted by questioning the rationale for the study locations previously selected and by the prospect of setting this concern aside with analyses conducted using a defined sampling frame. Additionally, advances in hardware and software made this type of analysis feasible. The NMMAPS project was initiated at the end of 1996, as PEEP was ending.

NMMAPS PART I: METHODS AND METHODOLOGIC ISSUES

The five sections of Part I address the areas of methodologic concern to NMMAPS.

- Exposure measurement error (sections 1 and 2) offers an overall conceptual framework for considering this methodologic problem and an approach for evaluating and correcting for bias from measurement error.
- Mortality displacement (sections 3 and 4) offers 2 analytic approaches for assessing the pollution-mortality relationship on different timeframes, including longer time scales that should be harvesting or mortality displacement–resistant.
- Methods for multicity analyses of air pollution and mortality (section 5), based on a new hierarchical approach for combining evidence from multiple locations.

The application of this method to the 20 largest cities in the US is provided only to demonstrate the method. The Part II report provides findings from detailed analysis of mortality data from these 20 cities and an exploration of heterogeneity in the effect of air pollution on mortality across the largest 90 cities. The second report also provides findings from analysis of data on hospitalization of individuals 65 years and older in the files of the HCFA.

EXPOSURE MEASUREMENT ERROR

Observational studies are inherently subject to bias from misclassification of outcome measures and exposure variables, whether the exposures are those directly under investigation or potential confounding or modifying factors. Although epidemiologic terminology generally refers to *information bias* or *misclassification,* common statistical parlance is *measurement error,* a term that has been used most often in discussing exposure misclassification in time-series studies of air pollution. This report uses either the epidemiologic or statistical terminology, as appropriate. In studies at the individual level, information bias may have many sources, including faulty recollection of past events or other misreporting by participants; problems with interviewing techniques or data collection procedures; and inaccuracy of measuring equipment (such as spirometers to measure lung function). There is an extensive literature on information bias in studies conducted at the individual level. Two reviews of this literature are by Armstrong and colleagues (1992) and Rothman and Greenland (1998).

Time-series studies of air pollution and mortality and morbidity are conducted on groups rather than individuals, a design generally termed an *ecological study*. The consequences of misclassification in ecological studies have received less attention because this design is most commonly used for generating hypotheses rather than hypothesis testing and risk estimation (Brenner et al 1992; Rothman and Greenland 1998). Ecological studies, however, have advantages over individual-level studies with respect to measurement error.

Consideration of the design assumptions of the time-series studies indicates the potential for their results to be affected by information bias. In these studies, measurements of concentrations taken at monitors generally sited for regulatory purposes are used as surrogates for exposures of those persons within the population at risk for morbidity or mortality. Even the most basic microenvironmental model, however, would show that personal exposures to pollutants such as particles depend on the specific environments where time is spent, the amount of time spent in these locations, and the concentrations of pollutant(s) in these environments (National Research Council [NRC] 1991). We also know that little time is spent outdoors on average and, consequently, indoor microenvironments contribute substantially to personal exposure for some pollutants; exposures indoors from outdoor pollutants that have penetrated into indoor spaces are relevant.

When time-series studies on mortality were reported, exposure misclassification was offered as a severe limitation in interpreting the observational evidence (Lipfert and Wyzga 1995, 1997). The potential error from using ambient concentrations measured for regulatory purposes as a surrogate for personal exposure was highlighted. Additionally, the potential for measurement error to cloud interpretation of multipollutant regression models was raised because varying degrees of measurement error for different pollutants might bias effect estimates. In the multipollutant case,

the consequences for interpreting findings for particulate matter (PM) indices depend on the relative degrees of measurement error for the pollutants considered in a model and on the pattern of correlation among the pollutants. The publications cited here and related discussions of the measurement error problem, although appropriately heightening awareness of this potential problem in the time-series studies, were not solidly grounded in statistical and epidemiologic theory. Because of the intense focus on these problems at the time NMMAPS was initiated, one of the specific objectives of the project was to address measurement error and its consequences in time-series studies of air pollution. To this end, the NMMAPS investigators, with input from members of the HEI-appointed Oversight Committee, which included members of the HEI Research Committee and other experts, formulated a conceptual framework for the measurement error problem and discussed this framework at a small workshop held in July 1997. These discussions evolved into a comprehensive report on measurement error in time-series studies of air pollution, which appears as section 1 of this Research Report, "Exposure Measurement Error in Time-Series Studies of Air Pollution."

Section 1 provides an overview of the main ideas on exposure measurement error in regression models. There are 2 types of errors—Berkson and classical—each with differing implications for interpreting model findings. In the classical error model, the expectation of the measured value is the true value, but there is variation—the measurement error—around the true value. In the Berkson error model, the expectation of the measured values is not the true value, but the average value of the true values for persons within each stratum of the measured value. The consequences of measurement error in univariate and multivariate models are also addressed. In the conceptual framework offered in section 1, 3 components of error are identified: deviation of individual personal exposure from the risk-weighted average personal exposure; the difference between the average personal exposure and the true ambient level; and the difference between the measured and the true ambient level. The analysis suggests that the second component, a classical error, is likely to be the most important source of bias. This framework can be used as a basis for adjusting for the consequences of measurement error by using data that provide an estimate of the bias associated with a particular error component.

This possibility of using data from an exposure assessment study to estimate and adjust for bias from measurement error was illustrated using data of the Particle Total Exposure Assessment Methodology (PTEAM) study (Ozkaynak et al 1996). PTEAM involved measurement of personal exposure to PM_{10} for 178 residents of Riverside, California. Ambient concentration data were also available so that the magnitude of the second component of error could be estimated. The PTEAM data were then used in a measurement error correction model with mortality data for Riverside.

This work on exposure measurement error was extended in section 2 of this Research Report, "A Measurement Error Model for Time-Series Studies of Air Pollution and Mortality." Additional data sets including personal exposure measurements and ambient concentrations were obtained from the original investigators. An enhanced 2-stage measurement-error correction model was developed that was a combination of (1) a Bayesian hierarchical generalized additive model and errors-in-variables for the mortality-ambient concentration relationship, and (2) a Berkson model for the ambient concentration–personal exposure relationship. This model is also applied as an illustration to the time series of mortality data from Baltimore.

Together, these 2 sections advance the formulation of the measurement error problem in time-series studies of air pollution and offer a measurement error correction model. These formal analyses of the problem of measurement error and its implications for interpreting the time-series studies show that measurement error is not likely to be as severe a limitation as proposed. In fact, we conclude that the generic criticism—that measurement errors render the result of time-series models to be uninterpretable—is incorrect. We have identified 3 measurement error components; our analysis suggests that the largest biases in inferences about the mortality relative risk occur due to the classical error component arising from using ambient concentration to estimate risk-weighted average personal exposures.

The framework guided the development of an approach for correcting for measurement error. The critical measurement error component can be estimated by comparing ambient concentrations to measured personal exposures. Fortunately, some data sets are available that contain the needed measurements. The hierarchical model proposed in section 2 offers a statistical tool for correcting for measurement error using data from personal exposure studies. The utility of this measurement error model is demonstrated by application to mortality data for Baltimore, Maryland, in which measurement error tended to blunt the exposure-response relationship.

Formal conceptualization of the measurement error problem clarified the data that should be collected for use in measurement error correction models. We found only limited data that could be used in the measurement error model; careful searching uncovered only 5 data sets for further analysis. Although a number of studies are now in progress on the relation between personal exposure and

ambient concentration, most do not have the longitudinal element needed for measurement error models. The data requirements of the measurement error models should be considered when exposure assessment studies are designed. Our example considered measurement error for a single pollutant because personal and ambient observations that covered multiple pollutants were not available. The methods developed and applied here could extend naturally to the multipollutant case.

The component of NMMAPS directed at measurement error was motivated by the need to characterize the consequences of exposure measurement error in time-series analysis and also to develop a systematic framework for adjusting for measurement error. The regulations of the EPA are directed at ambient concentrations and not personal exposures. The initial finding that the adjusted effect using the PTEAM data exceeds the estimate based on ambient monitoring data does not have immediate regulatory implications. However, varying degrees of measurement error may contribute to heterogeneity of effects across cities in the United States. We also need to explore the relations between ambient concentrations and personal exposures for groups considered susceptible to particulate matter.

MORTALITY DISPLACEMENT

The time-series studies of air pollution and mortality have shown associations between daily death counts, particularly from cardiac and respiratory causes, and levels of air pollution on the same or recent days. These associations have been widely interpreted as reflecting the effect of air pollution on a group of individuals who have heightened susceptibility because of their chronic heart or lung diseases. A simple 2-compartment model that divides the population into a frail group at risk for dying and the remaining nonfrail persons can represent this characterization of the population. In this simple model, depletion of the frail pool would reduce the subsequent number of deaths.

For interpreting the observed associations under the assumption of a frailty model, an understanding of the degree of life shortening underlying the associations is needed. At the extreme, if the effect of air pollution were to advance the timing of death by only a brief interval, such as a single day, then the net loss of life would be limited and, in fact, the time lost might arguably be a time characterized by a low quality of life for the frail individuals at risk for dying. The unfortunate term *harvesting* has been applied to describe this phenomenon; that is, only extremely frail individuals die from air pollution, sustaining only a slight reduction of their life spans (Schimmel and Murawski 1976). To avoid this somewhat objectionable label, the phrase *mortality displacement* has also been used and is the phrase used in this Research Report. This phenomenon was raised as a limitation of the findings of daily time-series studies (Lipfert and Wyzga 1995) and was directly discussed in the Staff Paper on Particulate Matter prepared by the EPA (1996). The findings of the long-term prospective cohort studies of air pollution and mortality—the Harvard Six Cities Study (Dockery et al 1993) and the American Cancer Society's Cancer Prevention Study (CPS) II (Pope et al 1995)—were considered to offer critical evidence counter to the mortality displacement hypothesis.

Other investigators have approached the problem of mortality displacement. If the association between air pollution and mortality does reflect the existence of a pool of frail individuals in the population, then episodes of high pollution that lead to increased mortality might reduce the size of this pool and subsequent days would then be expected to show a reduced effect of air pollution. The occurrence of this phenomenon could be assessed by testing for interaction between prior high-pollution days and subsequent pollution exposure on mortality counts; under the mortality displacement hypothesis, a negative interaction is predicted. Spix and colleagues (1993) analyzed the daily time series of deaths in Erfurt in the former East Germany for the period 1980 to 1988. Weak evidence for this interaction was found. Drs L Cifuentes and LB Lave (Carnegie-Mellon University, Pittsburgh PA, unpublished data, 1996) also assumed a frailty model and proposed that episodes of air pollution involving a high level of pollution followed by a low level would have a profile of mortality that was initially high and was then below expectation because of the depletion of susceptible individuals. They also found evidence for mortality displacement, applying this approach to data for Philadelphia for 1983 to 1988. Spix and colleagues (1993) proposed a 1-step Markov chain model and demonstrated that the pollution relative risk estimates from Poisson regression are biased about 10% to 30% by mortality displacement. Smith and colleagues (1997) used a 2-compartment model with the additional assumption that both the risk of becoming frail and the risk of death may depend on air pollution.

To date, these methods have had limited application and the degree of mortality displacement in the effects estimated in the daily time-series studies remains unresolved. In NMMAPS, a conceptually different approach that uses the daily time-series data to assess associations on short and long time scales was proposed. Two closely related methods for analysis of daily time-series data were developed, both testing for air pollution–mortality associations on varying time scales. The approaches are unified by the

underlying concept that mortality displacement should introduce association on shorter time scales while longer time scales should be mortality-displacement–resistant. In section 3, "Mortality Displacement–Resistant Estimates of Air Pollution Effects on Mortality," Zeger and colleagues conducted a simulation study using 3 time series of mortality generated with differing underlying assumptions about the length of residence time spent in the frail state—3, 30, and 300 days—but assuming the same relative risk from pollution. The correlation between the number of deaths and the particle index became nonnegligible only at time scales less than about twice the mean residence time. Thus, mortality displacement introduces association between mortality and pollution only at shorter time scales.

In the approach developed by Zeger, Dominici, and Samet and described in section 3, frequency domain log-linear regression (Kelsall et al 1997) was used to decompose the information about the pollution-mortality association into distinct time scales. This technique, developed and applied to air pollution and mortality as a part of PEEP, decomposes the air pollution series and the mortality series into distinct component series so that associations between pollution and mortality can be calculated for each time scale. Under the mortality displacement hypothesis, we expect to find an association only on short time scales if mortality displacement is the only cause of the pollution-mortality association.

As described in section 4, "Mortality Displacement and Long-Term Exposure Effects Related to Air Pollution and Mortality," Schwartz used Cleveland's STL (seasonal and trend decomposition using LOESS [locally weighted smoother]) filtering algorithm to separate the time series of daily deaths, air pollution, and weather into long wavelength components, midscale components, and residual very short time scale components. Associations in the long wavelength components of the data were assumed to represent effects of time trends and seasonal fluctuations, and short-term components were assumed to reflect short-term mortality displacement. The long-term component was set aside by using a LOESS smooth, a weighted moving regression, with a window of 120 days. The remaining, midscale information was used to assess the association between air pollution and mortality without confounding from longer-term effects or liability to detecting mortality displacement. The midscale components were examined with smoothing windows of 15, 30, 45, and 60 days.

In this report, the 2 methods have the same underlying analytic principle: decomposition of the daily time-series data into its shorter- and longer-term components. Zeger and colleagues (section 3) apply frequency domain regression to data for Philadelphia, 1974–1988, while Schwartz (section 4) uses data for Boston, 1979–1986. These new methods complement prior approaches. In this report, we present these methods and illustrate their application.

In NMMAPS, we have developed 2 conceptually similar, albeit computationally distinct, approaches for testing for effects of air pollution on mortality on time scales not affected by mortality displacement. Zeger and colleagues show, using a simulated example involving a simple 2-compartment frailty state model, that mortality displacement does not produce associations on time scales longer than about twice the mean residence time in the frail pool. Consequently, effects of air pollution on mortality that are demonstrated on longer time scales should not reflect mortality displacement alone. The frequency domain regression approach used by Zeger and colleagues involves a continuous decomposition of the data into distinct time axes. Schwartz uses a filtering algorithm, the STL algorithm, to separate the time-series information into long, mid, and very short time-scale components.

The two methods are closely related. The Zeger and colleagues method gives a continuous smooth estimate of relative risk as a function of time scale. It can be averaged over distinct ranges of the time-scale to produce the long, mid, and short time-scale components used by Schwartz. Alternately, Zeger and colleagues' relative risk function can be averaged over a predetermined range of longer time scales, omitting the shorter scales where mortality displacement will create associations, to obtain the displacement-resistant estimator of pollution effects. The Schwartz approach has the advantage of being easy to implement with standard software. Both of the approaches discussed here are distinct from the approaches of Spix and associates (1993) and Smith and coworkers (1997), who attempt to make inferences about the size of the frail population using parametric latent variable models.

Zeger and colleagues apply their method to the previously analyzed data for Philadelphia, 1973 to 1988 (Samet et al 1997). They implement frequency-domain linear regression, while adjusting for temperature and dew point and longer-term trends following the approach of the earlier analysis. In the Philadelphia data, they find little evidence for mortality displacement in assessing the effect of particulate matter less than 10 μg in aerodynamic diameter (PM_{10}). In fact, the association is present on longer time scales and tends to diminish at the shorter time scales on which mortality displacement would operate.

In section 4, Schwartz applies his method to data for the city of Boston from 1979 to 1986. The pattern of midscale associations varied by cause of death. The pattern of variation

of the 4 estimates of risk at the various smoothing frequencies was taken as an indication of the presence of mortality displacement. An effect that diminished from the shorter-term to the longer-term scales was interpreted as evidence for mortality displacement, while persistent or increasing effects at the longer-term scales were interpreted as an effect beyond mortality displacement. The pattern for chronic obstructive pulmonary disease (COPD) was consistent with mortality displacement; the pattern for pneumonia was consistent with both some mortality displacement and a longer-term effect; and the patterns for ischemic heart disease and total mortality were consistent with longer-term effects without evidence for mortality displacement. The equivalent pattern for total mortality was observed with frequency domain regression in the analysis of Philadelphia data reported in section 3.

These 2 sections demonstrate analytic methods that can be used to detect the presence of mortality displacement and to gauge the extent to which associations found with conventional time-series methods reflect mortality displacement versus longer-term effects. Evidence was found in Philadelphia and in Boston that associations of particle indices with mortality did not represent mortality displacement alone. The next steps include applying these methods to other locations and developing a 2-stage analog of frequency domain regression to assess mortality displacement more powerfully by combining evidence from multiple locations. Additionally, more formal comparisons of the two methods described in this report, to each other and to other approaches, are needed to determine the most informative approaches for assessing mortality displacement.

MULTICITY MODELS: COMBINING EVIDENCE FROM MULTIPLE LOCATIONS

The evidence on particulate air pollution and mortality that initially renewed public health concern about the adverse health effects of air pollution came from time-series analyses of data from single cities. The locations had been selected primarily on the basis of historical precedent as in London (Schwartz and Marcus 1990), data availability as in Philadelphia (Schwartz and Dockery 1992), or pollution or source characteristics as in Santa Clara (Fairley 1990). The generality of evidence from these locations was uncertain, but subsequent studies were consistent with the initial reports in spite of variation in the methods of the studies (Dockery and Pope 1994). Little evidence of heterogeneity was found with meta-analysis of effect estimates in studies published through 1994 (Dockery and Pope 1994; Schwartz 1994), and summary estimates for the effect of particulate matter on mortality were statistically significant. Nonetheless, critics continued to question the interpretation that this evidence was indicative of an effect of particles specifically rather than of air pollution generally (Moolgavkar and Luebeck 1996; McClellan 1997).

Concern about the representative nature of findings in particular locations can be addressed by selecting multiple study locations from a defined sampling frame. This approach was followed in NMMAPS, which used all cities in the United States with PM_{10} monitors as the sampling frame. This design set aside concern as to the representative nature of the study locations. The design also brought heterogeneity in the levels of pollutants other than PM_{10} and thereby facilitated exploration of the independence of the PM_{10} effect.

This approach was anticipated in the multicenter European study, Air Pollution and Health: A European Approach (APHEA) (Katsouyanni et al 1997; Touloumi et al 1997). In APHEA, 12 cities were selected from both western and central Europe, although not on a systematic basis. Data on particulate air pollution and daily mortality were analyzed according to a standardized protocol. Model estimates from the individual cities were pooled as the weighted means of the regression coefficients, and heterogeneity among cities was explored using a random effects model.

In NMMAPS, we developed hierarchical regression models for combining estimates of the pollution-mortality relation across cities. The analysis proceeded in 2 stages. Given a time series of daily mortality counts in each of 3 age groups, we used generalized additive models (Hastie and Tibshirani 1990) to estimate the relative change in the rate of mortality associated with changes in the air pollution variables, controlling for age-specific longer-term trends, weather, and other potential confounding factors, separately for each city. We then combined the pollution mortality relative rates across the cities using a Bayesian hierarchical model (Lindley and Smith 1972; Morris and Normand 1992). This approach allowed us to obtain an overall estimate and to explore whether some of the geographic variation can be explained by site-specific explanatory variables.

In section 5, "Combining Evidence on Air Pollution and Daily Mortality from Twenty Largest US Cities," we offer a comprehensive description of the analytic approach and demonstrate its application. NMMAPS Part II will provide the findings from the proposed sampling frame of 90 cities.

CONCLUSIONS

NMMAPS represents a comprehensive research program on time-series analysis of data on air pollution and morbidity and mortality. Its elements include the development

of a conceptual framework for the problem of measurement error and of a related measurement error correction model, consideration of informative analytic approaches to assessing the extent of mortality displacement, and the development of analytic methods for carrying out multicity analyses. These elements are the subject of this report; the second report will cover the multicity morbidity analyses.

To summarize briefly, the evaluation of measurement error and its consequences shows that the critical component of error would tend to blunt effects; if appropriate personal exposure data are available, a correction for bias can be made. Two conceptually linked methods for addressing mortality displacement were developed and applied. Findings in Philadelphia and Boston indicated associations on longer, mortality displacement–resistant time scales. A new method for multicity analysis was developed and its potential shown by analysis of data for 20 US cities.

The findings in this report indicate a number of informative applications and extensions of the NMMAPS methodology:

- optimization of the design of exposure assessment studies for measurement error correction;
- application of the frequency domain regression method and the STL algorithms to data from additional cities; and
- periodic multicity analyses for surveillance of changing effects of air pollution.

REFERENCES

Armstrong BK, Saracci R, White E. 1992. Principles of Exposure Measurement in Epidemiology. Oxford University Press, New York NY.

Brenner H, Savitz DA, Jockel KH, Greenland S. 1992. The effects of nondifferential exposure misclassification in ecological studies. Am J Epidemiol 135:85–95.

Dockery DW, Pope CA III. 1994. Acute respiratory effects of particulate air pollution. Annu Rev Public Health 15:107–132.

Dockery DW, Pope CA III, Xu X, Spengler JD, Ware JH, Fay ME, Ferris BG Jr, Speizer FE. 1993. An association between air pollution and mortality in six US cities. N Engl J Med 329:1753–1759.

Fairley D. 1990. The relationship of daily mortality to suspended particulates in Santa Clara County, 1980–1986. Environ Health Perspect 89:159–168.

Hastie TJ, Tibshirani RJ, eds. 1990. Generalized Additive Models. Chapman & Hall, New York NY.

Katsouyanni K, Touloumi G, Spix C, Schwartz J, Balducci F, Medina S, Rossi G, Wojtyniak B, Sunyer J, Bacharova L, Schouten JP, Ponka A, Anderson HR. 1997. Short-term effects of ambient sulphur dioxide and particulate matter on mortality in 12 European cities: Results from the APHEA project. Br Med J 314:1658–1663.

Kelsall JE, Samet JM, Zeger SL, Xu J. 1997. Air pollution and mortality in Philadelphia, 1974–1988. Am J Epidemiol 146:750–762.

Lindley DV, Smith AFM. 1972. Bayes estimates for the linear model (with discussion). J R Stat Soc 34:1–41.

Lipfert FW, Wyzga RE. 1995. Air pollution and mortality: Issues and uncertainties. J Air Waste Manage Assoc 45:949–966.

Lipfert FW, Wyzga RE. 1997. Air pollution and mortality: The implications of uncertainties in regression modeling and exposure measurement. J Air Waste Manage Assoc 47:517–523.

McClellan RO, Miller FJ. 1997. An overview of EPA's proposed revision of the particulate matter standard. CIIT Activities 17:1–24.

Moolgavkar SH, Luebeck EG. 1996. A critical review of the evidence on particulate air pollution and mortality. Epidemiology 7:420–428.

Morris CN, Normand S-L. 1992. Hierarchical models for combining information and for meta-analysis. In: Bayesian Statistics 4. Proceedings of the Fourth Valencia International Meeting (Bernardo JM, Berger JO, Dawid AP, Smith AFM, eds). Oxford University Press, Oxford, England.

National Research Council. 1991. Committee on Advances in Assessing Human Exposure to Airborne Pollutants. Human Exposure Assessment for Airborne Pollutants: Advances and Opportunities. National Academy Press, Washington DC.

Ozkaynak H, Xue J, Spengler J, Wallace L, Pellizzari E, Jenkins P. 1996. Personal exposure to airborne particles and metals: Results from the Particle TEAM Study in Riverside, California. J Expo Anal Environ Epidemiol 6:57–78.

Pope CA, III, Thun MJ, Namboodiri MM, Dockery DW, Evans JS, Speizer FE, Heath CW Jr. 1995. Particulate air pollution as a predictor of mortality in a prospective study of US adults. Am J Respir Crit Care Med 151:669–674.

Rothman KJ, Greenland S. 1998. Modern Epidemiology, 2nd ed. Lippincott-Raven, Philadelphia PA.

Samet JM, Zeger S, Dominici F, Curriero F, Coursac I, Dockery D, Schwartz J, Zanobetti A. 2000. The National Morbidity and Mortality Air Pollution Study. Part II. Morbidity and Mortality from Air Pollution in the United States. Research Report 94. Health Effects Institute, Cambridge MA.

Samet JM, Zeger S, Kelsall JE, Xu J, Kalkstein LS. 1998. Does weather confound or modify the association of particulate air pollution with mortality? An analysis of the Philadelphia data, 1973–1980. Environ Res 77:9–19.

Samet JM, Zeger SL, Berhane K. 1995. The association of mortality and particulate air pollution. In: Particulate Air Pollution and Daily Mortality: Replication and Validation of Selected Studies, The Phase I.A Report of the Particle Epidemiology Evaluation Project. Health Effects Institute, Cambridge MA.

Samet JM, Zeger SL, Kelsall JE, Xu J, Kalkstein LS. 1997. Air pollution, weather, and mortality in Philadelphia 1973–1988. In: Particulate Air Pollution and Daily Mortality: Analyses of the Effects of Weather and Multiple Pollutants. The Phase I.B Report of the Particle Epidemiology Evaluation Project. Health Effects Institute, Cambridge MA.

Schimmel H, Murawski TJ. 1976. The relation of air pollution to mortality. J Occup Med 18:316–333.

Schwartz J. 1994. Air pollution and daily mortality: A review and meta-analysis. Environ Res 64:36–52.

Schwartz J, Dockery DW. 1992. Increased mortality in Philadelphia associated with daily air pollution concentrations. Am Rev Respir Dis 145:600–604.

Schwartz J, Marcus A. 1990. Mortality and air pollution in London: A time-series analysis. Am J Epidemiol 131:185–194.

Smith L, Davis J, Sacks J, Speckman P, Styer P. 1997. Assessing the human health risk of atmospheric particles. In: Statistics and the Environment. Proceedings from the 1997 Joint Statistical Meetings, Anaheim CA. American Statistical Association, Alexandria VA.

Spix C, Heinrich J, Dockery DW, Schwartz J, Volksch G, Schwinkowski K, Collen C, Wichmann HE. 1993. Air pollution and daily mortality in Erfurt, East Germany, 1980–1989. Environ Health Perspect 101:518–526.

Touloumi G, Katsouyanni K, Zmirou D, Schwartz J, Spix C, Ponce de Leon A, Tobias A, Quenel P, Rabczenko D, Bacharova L, Bisanti L, Vonk JM, Ponka A. 1997. Short-term effects of ambient oxidant exposure on mortality: A combined analysis within the APHEA project. Am J Epidemiol 146:177–185.

US Environmental Protection Agency. 1996. Review of the National Ambient Air Quality Standards for Particulate Matter: Policy Assessment of Scientific and Technical Information. OAQPS Staff Paper. EPA-452\R-96-013. Office of Air Quality Planning and Standards, Research Triangle Park NC.

Zeger SL, Dominici F, Samet J. 1999. Harvesting-resistant estimates of air pollution effects on mortality. Epidemiology 10:171–175.

ABOUT THE PRIMARY AUTHORS

Jonathan M Samet is Professor and Chairman of the Department of Epidemiology at Johns Hopkins School of Public Health. He received his medical degree from the University of Rochester School of Medicine and Dentistry and a master's degree in epidemiology from Harvard School of Public Health. His research has focused on the effects of environmental and occupational agents. He is chairman of the NRC's Committee on Research Priorities for Airborne Particulate Matter and was chairman of the NRC's Committee on Health Risk of Exposure to Radon (BEIR VI).

Francesca Dominici, Assistant Professor of Biostatistics at the Johns Hopkins School of Public Health, received her doctoral degree from the University of Padua. Her research interests include time-series analysis in environmental epidemiology, semiparametric Bayesian models for dose-response data, hierarchical models for combining information and for meta-analysis, and missing data. She is the first recipient of the Walter A Rosenblith New Investigator Award from the Health Effects Institute. Her work on air pollution and mortality was recently honored by an invitation to read her paper on this topic to the Royal Statistical Society in London.

Scott L Zeger is Professor and Chairman of the Department of Biostatistics at Johns Hopkins School of Public Health. He earned his doctorate in statistics from Princeton University. With Kung-Yee Liang, he developed the generalized estimating equation (GEE) approach to regression modeling with correlated data as occur in time-series or longitudinal studies. This work has been recognized with awards from both the American Statistical Association and American Public Health Association. Zeger's current research focuses

on improving the design and analysis of data from prospective studies.

Joel Schwartz, Associate Professor of Environmental Epidemiology at Harvard School of Public Health and Harvard Medical School, received his doctorate from Brandeis University in theoretical physics and is a MacArthur Fellow. Schwartz has held appointments as a visiting scientist at the University of Basel, Switzerland, and at the University of Wupperal, Germany, and has served as a Senior Scientist at the US EPA. His research focuses on the health effects of air and water pollution, lead, and methodologic questions regarding the modeling of covariates in epidemiologic studies. Over the past decade, his studies have been instrumental in precipitating renewed interest in airborne particles.

Douglas W Dockery is Professor of Environmental Epidemiology at Harvard School of Public Health and Professor of Epidemiology at Harvard Medical School. He received his master's degree in meteorology from the Massachusetts Institute of Technology and his doctorate from Harvard School of Public Health in environmental health sciences. His research interests focus on the health effects of air pollution, asthma, lead toxicity, and water quality. He is currently president of the International Society of Environmental Epidemiology.

PUBLICATIONS RESULTING FROM THIS RESEARCH

Dominici F, Samet J, Xu J, Zeger S. 2000. Combining evidence on air pollution and daily mortality from the 20 largest US cities: A hierachical modeling strategy. J R Stat Soc, Series C. In press.

Dominici F, Zeger SL, Samet JM. 2000. A measurement error model for time-series studies of air pollution and mortality. Biostatistics. In press.

Kelsall JE, Samet JM, Zeger SL, Xu J. 1997. Air pollution and mortality in Philadelphia, 1974–1988. Am J Epidemiol 146:750–762.

Kelsall J, Zeger S, Samet J. 1999. Frequency domain log-linear models: Air pollution and mortality. Appl Stat 48 (Pt 3):331–334.

Samet JM, Zeger SL, Kelsall JE, Xu J. 1998. Particulate air pollution and mortality: The Particle Epidemiology Evaluation Project. Appl Occup Environ Hyg 13:364–369.

Samet JM, Zeger SL, Kelsall JE, Xu J, Kalkstein LS. 1998. Does weather confound or modify association of particulate air pollution with mortality? An analysis of the Philadelphia data, 1973–1980. Environ Res 77:9–19.

Samet JM, Zeger S, Dominici F, Curriero F, Coursac I, Dockery D, Schwartz J, Zanobetti A. 2000. The National Morbidity and Mortality Air Pollution Study. Part II. Morbidity and Mortality from Air Pollution in the United States. Research Report 94. Health Effects Institute, Cambridge MA.

Schwartz J. 2000. Harvesting and long-term exposure effects in the relationship between air pollution and mortality. Am J Epidemiol 151:440–448.

Zeger S, Thomas D, Dominici F, Cohen A, Schwartz J, Dockery D, Samet J. 2000. Exposure measurement error in time-series studies of air pollution: Concepts and consequences. Environ Health Perspect. In press.

Zeger SL, Dominici F, Samet J. 1999. Harvesting-resistant estimates of air pollution effects on mortality. Epidemiology 10:171–175.

ABBREVIATIONS AND OTHER TERMS

AIRS	Aerometric Information Retrieval System
APHEA	Air Pollution and Health: A European Approach
CI	confidence interval
CO	carbon monoxide
CM	coarse mass
CODA	Convergence Diagnostics and Output Analysis (software)
COPD	chronic obstructive pulmonary disease
df	degrees of freedom
EPA	US Environmental Protection Agency
FDLLR	frequency domain log-linear regression
HCFA	Health Care Financing Administration
ICD-9	International Classification of Diseases, Ninth Revision
IHD	ischemic heart disease
IQR	interquartile range
LOESS	locally weighted smoother
NMMAPS	National Morbidity, Mortality, and Air Pollution Study
MCMC	Markov chain Monte Carlo
MRT	mean residence time
NAAQS	National Ambient Air Quality Standard

NCHS	National Center for Health Statistics
NETH-A	Netherlands study of adults
NETH-C	Netherlands study of children
NO_2	nitrogen dioxide
NRC	National Research Council
O_3	ozone
PEEP	Particle Epidemiology Evaluation Project
PM	particulate matter
$PM_{2.5}$	particulate matter less than 2.5 μm in aerodynamic diameter
PM_{10}	particulate matter less than 10 μm in aerodynamic diameter
PTEAM	Particle Total Exposure Assessment Methodology (study)
SO_2	sulfur dioxide
STL	seasonal and trend decomposition using LOESS
THEES	Total Human Environmental Exposure Study
TSP	total suspended particles

Section 1: Exposure Measurement Error in Time-Series Studies of Air Pollution

Scott L Zeger, Duncan Thomas, Francesca Dominici, Jonathan M Samet, Joel Schwartz, Douglas W Dockery, and Aaron Cohen

ABSTRACT

Misclassification of exposure has long been recognized as an inherent limitation of epidemiologic studies of the environment and disease. For many agents of interest, exposures take place over time and in multiple locations; accurately estimating the relevant exposures for an individual participant in epidemiologic studies is often daunting, particularly within the limits set by feasibility, participant burden, and cost. The problem of measurement error is well recognized, and researchers have taken steps to deal with its consequences by limiting the degree of error through the design of a study, estimating the degree of error using a nested validation study, and making adjustments for measurement error in statistical analyses.

Section 1 sets out a systematic conceptual formulation of the problem of measurement error in epidemiologic studies of air pollution and considers the consequences of measurement error within this formulation. When possible, available data were used to make simple estimates of measurement error effects.

The introduction to section 1 presents an overview of the main ideas on measurement errors in linear regression, distinguishing 2 extremes of a continuum: Berkson from classical type errors, and the univariate predictor from the multivariate predictor case. We then propose a single conceptual framework for evaluation of measurement errors in the log-linear regression used for time-series studies of particulate air pollution and mortality, identifying 3 main components of error. We also present new simple analyses of data on exposures of particulate matter less than 10 µm in aerodynamic diameter (PM_{10})* from the Particle Total Exposure Assessment Methodology (PTEAM) study (Ozkaynak et al 1996). Finally, we summarize open questions regarding measurement error and suggest the kind of additional data necessary to address them.

INTRODUCTION

Misclassification of exposure has long been recognized as an inherent limitation of epidemiologic studies of the environment and disease (Armstrong et al 1992). For many agents of interest, exposures take place over time and in multiple locations so that it is difficult to estimate accurately the relevant exposures for individual study participants, particularly within the limits set by feasibility, participant burden, and cost. In general, exposure measurement error tends to blunt the sensitivity of epidemiologic studies for detecting effects of environmental agents. The specific impact of exposure error on effect estimates depends on several factors including the study design, the types of error, and the relationships between outcome and independent variables (Armstrong et al 1992; Thomas et al 1993). As the problem of exposure error has become recognized, researchers have taken steps to control its consequences by limiting the degree of error through careful study design and data collection, by estimating the degree of error using a nested validation study, and by making adjustments for measurement error in statistical analyses.

This section addresses the problem of exposure error in observational, ecological time-series studies of air pollution and health. Pollution of outdoor air is a public health concern throughout the world. For decades, epidemiologic studies have been a cornerstone to investigating the health effects of air pollution and have been a principal basis for setting regulations to protect the public against adverse health effects. The 2 broad types of observational study designs are ecological or aggregate-level studies (either cross-sectional or time-series design) and individual-level studies (primarily cross-sectional or cohort designs). In

The National Morbidity, Mortality and Air Pollution Study: Methods and Methodologic Issues, Part I of Health Effects Institute Research Report 94, includes an Investigators' Report, a Preface, a Commentary by the Health Review Committee, and an HEI Statement about the research project. Correspondence concerning this section may be addressed to Dr Scott L Zeger, Johns Hopkins School of Public Health, Department of Biostatistics, 615 North Wolfe Street, Baltimore MD 21205-2179.

Although this document was produced with partial funding by the United States Environmental Protection Agency under Assistance Award R824835 to the Health Effects Institute, it has not been subjected to the Agency's peer and administrative review and therefore may not necessarily reflect the views of the Agency, and no official endorsement by it should be inferred. The contents of this document also have not been reviewed by private party institutions, including those that support the Health Effects Institute; therefore, it may not reflect the views or policies of these parties, and no endorsement by them should be inferred.

* A list of abbreviations and other terms appears on page 13.

ecological studies, population-level indicators of exposure are typically drawn from centrally sited air pollution monitors. In individual-level cross-sectional and cohort studies, exposure estimates for individual participants may be based on centrally located monitors, on a combination of central monitors with personal records of environments where participants spend time, or on personal exposure monitoring (National Research Council [NRC] 1991).

Regardless of study design, any pollution exposure assessment strategy introduces some degree of exposure measurement error. For example, in the Harvard Six Cities study, a prospective cohort assessment of air pollution and respiratory health and mortality, exposure estimates for persons from each of the 6 cities were based on centrally sited monitors (Ferris et al 1979; Dockery et al 1993). Exposures were further characterized for samples of participants by using personal monitors and monitors placed in their homes; the resulting data provide an understanding of the components of error associated with using central site data for all participants.

The problem of measurement errors in predictor variables in regression analysis has been carefully studied in the statistics and epidemiologic literature for several decades. Fuller (1987) summarizes early research on linear regression with "errors-in-x" variables. Carroll and colleagues (1995) extend this literature to generalized linear models including Poisson, logistic, and survival regression analyses. Thomas and colleagues (1993) present an overview of the exposure error or "misclassification problem" from the general epidemiologic perspective. For recent illustrations of statistical approaches to measurement error in epidemiologic research, see Spiegelman and colleagues (1997), Willett (1998), and Pierce and colleagues (1990).

In one of the early papers on exposure error in studies of air pollution, Shy and colleagues (1978) described the problem and addressed its consequences in an epidemiologic framework. Goldstein and Landovitz (1997a,b) recognized that a single monitoring station may not represent a geographic area adequately and conducted an analysis of correlations among concentration data from several monitors in New York. In the ensuing decades, there has been a deepening understanding of measurement error in general and its potential implications specifically for the study of air pollution (NRC 1985; Navidi et al 1994).

During the 1990s, substantial new evidence—largely from ecologic, time-series analyses of air pollution and mortality—showed that daily variation in ambient measures of particulate air pollution, within current standards of the US Environmental Protection Agency (EPA), was associated with daily mortality levels (Dockery and Pope 1994). Strong concerns have been raised about interpreting these associations in view of potential errors in the exposure measurements. In a series of papers, Lipfert (1997a,b) and Lipfert and Wyzga (1997) have suggested that the central monitoring data used in the time-series analyses have an uncertain relationship with exposures of individuals in the study communities; they have further argued that those errors vary among pollutants, complicating interpretation of any multipollutant models. Lipfert and Wyzga have referred specifically to an analysis by Schwartz, Dockery, and Neas (1996) that attributed effects on mortality to fine rather than coarse particles, based in part on the results of multivariable models that included variables for both particulate measures.

A number of exposure assessment studies have found sizable differences between actual personal exposures to particles and estimates based on central monitor values (eg, Wallace 1996). Some have questioned whether the observed associations are plausible given these findings. Schwartz, Dockery, and Neas have responded that, as the number of deaths per day is calculated over the population, the relevant exposure measure is the mean of personal exposures on that day, which is probably more tightly correlated with central station monitoring than individual exposures. Janssen and colleagues (1998) have reported that much of the variation in PM_{10} measurements is between people and that the longitudinal correlation between average and ambient PM_{10} measures is relatively high. The debate over measurement error and its consequences has taken place, however, without the development of a comprehensive formulation of the problem.

Because exposure measurement error may have substantial implications for interpreting epidemiologic studies on air pollution, particularly the time-series analyses, this section of the Investigators' Report describes one systematic conceptual formulation of the problem of exposure error in epidemiologic, time-series studies of air pollution and considers the possible consequences for relative risk estimation. We have used available and relevant data to obtain rough estimates of the magnitudes of the effects of measurement error for 1 city.

The next subsection presents an overview of the main ideas on exposure measurement errors in linear regression, distinguishing Berkson from classical type errors and the univariate from multivariate predictor cases. The conceptual framework for evaluation of measurement errors in the log-linear regression models used for time-series studies of particulate air pollution and mortality identifies 3 main components of error. We then present new analyses of data on exposures to PM_{10} from the PTEAM study (Ozkaynak et al 1996) and illustrate how data on personal and ambient exposure levels can be used to assess the effects of measurement

error on the estimated associations of PM_{10} with daily mortality. We also illustrate a statistical approach for assessing the bias in a relative risk estimate caused by exposure measurement error. We summarize the open questions regarding measurement error and propose the additional data needed to address these questions more effectively.

MEASUREMENT ERROR EFFECTS IN REGRESSION MODELS

The fundamental concepts of how exposure error can affect an epidemologic study of pollution and health are presented here by considering the effects of exposure measurement error in a standard linear Gaussian regression model. This topic has been treated in full detail elsewhere (Snedecor and Cochran 1980; Carroll et al 1984; Fuller 1987; Thomas et al 1993; Carroll et al 1995). For simplicity, consider a regression of the health response y_t (eg, log mortality rate on day t) and predictors x_t (eg, PM_{10}, O_3, weather, …):

$$y_t = \alpha + \beta_x x_t + \varepsilon_t \quad (1)$$

where α and β_x are regression coefficients to be estimated and ε_t represents residual error that is assumed to be independent of x_t. Here β_x is the expected change in mortality per unit change in true exposure. Given observations (x_t, y_t), $t = 1, \ldots, T$ and appropriate assumptions about the distribution of the residuals, ordinary least squares estimation provides optimal (unbiased and minimum varianced) estimates of the regression coefficients.

Now we assume that instead of the true exposure levels x_t, we have only an imperfect measure of exposure, denoted z_t. The overall difference between x_t and z_t comprises multiple components of error including differences: between individual- and population-average exposures; between population-average exposures and ambient levels at central sites; and between actual ambient levels and the measurements of those levels. Suppose we regress the health outcome y_t on the imperfect z_t rather than x_t, which is unavailable:

$$y_t = \alpha^* + \beta_z z_t + \varepsilon_t^*. \quad (2)$$

How will $\hat{\beta}_z$ differ from $\hat{\beta}_x$?

To answer this question, we will first assume that z_t is a surrogate for x_t (Carroll et al 1995), which means that, given x_t, there is no additional information in z_t about y_t. We then can distinguish 2 fundamentally distinct types of relationships between the true and measured exposures, which represent poles of a measurement error continuum. The first type is referred to as the *classical error model* (Carroll et al 1995) in which we assume that z is an imperfect measure of x, so that the average z within each x stratum equals x ($E(z \mid x) = x$). Then it follows that the measurement error, $(z - x)$, is uncorrelated with the true value x. This classical model is a reasonable one for the difference between measured ambient levels of pollution and the true values for a measuring device that is unbiased. That is, when the true level of pollution is x, an unbiased instrument will measure x on average, even if individual measurements z differ from x.

The second type of model for measurement error is the *Berkson error model* (Carroll et al 1995). In this model, we assume that the average value of the true exposure x within each stratum of measured level z equals z ($E(x \mid z) = z$). This Berkson model is appropriate when z represents a measurable environmental factor that is shared by a group of participants whose individual exposures x might vary because of time-activity patterns. For example, z might be the spatially averaged ambient level of a pollutant without major indoor sources and x might be the personal exposures, which, when averaged across people, match the ambient level.

Classical and Berkson models for exposure measurement errors represent 2 extremes of a continuum. Most exposure errors combine elements of each, but because the consequences for risk assessment of classical and Berkson errors differ, it is useful to consider each in turn. In the Berkson error case, if we regress y_t on z_t rather than on x_t, the estimate $\hat{\beta}_z$ is an unbiased estimate of the coefficient β_x, which would be obtained by regressing y_t on the actual exposure x_t. That is, having z_t rather than x_t does not lead to bias in the regression coefficients under the surrogacy assumption. The exposure measurement error does increase the variance of the regression coefficient, however, since having z_t rather than x_t is obviously not as informative about the coefficient β_x. Bias is not introduced, however. The same is true if the average x at each value of z differs from z by a fixed amount a, that is, $E(x \mid z) = z - a$.

In contrast, under the classical error model, $\hat{\beta}_z$ obtained by regressing y_t on the imperfect measure exposure z_t is a biased estimate of β_x. In the simple linear regression with 1 explanatory variable, $\hat{\beta}_z$ is expected to be smaller than β_x, or *attenuated*. The degree of attenuation increases as the variance of the exposure error increases. Again, a constant difference in the expected values of the 2 measures does not change this result.

It is useful to establish the results summarized above on the effects of exposure error on simple linear regression

coefficients and helpful to do so in advance of considering a multiple regression case. To reestablish notation, the model of interest is

$$y_t = \alpha + \beta_x x_t + \varepsilon_t \quad (1)$$

but because x_t is unobserved we instead might regress y_t on z_t:

$$y_t = \alpha^* + \beta_z z_t + \varepsilon_t^* \quad (2)$$

The question is how will $\hat{\beta}_z$ from (2) estimate β_x in (1). Under the Berkson error model ($E(x_t \mid z_t) = z_t$, so that we can also infer the regression of y on z from (1):

$$E(y_t \mid z_t) = \alpha^* + \beta_x E(x_t \mid z_t) = \alpha + \beta_x z_t. \quad (3)$$

Comparing (2) and (3) shows that in the Berkson error case: $\beta_z = \beta_x$; that is, $\hat{\beta}_z$ is an unbiased estimate of β_x. Adding a constant to 1 exposure variable affects only the intercept.

Under the classical model, z_t is assumed to vary about x_t or $E(z_t \mid x_t) = x_t$, which does not imply $E(x_t \mid z_t) = z_t$. If we further assume that x_t and $(z_t - x_t)$ are jointly normally distributed, it can be shown that

$$E(y_t \mid z_t) = \alpha^{**} + c\beta_x z_t$$

where c is an attenuation factor between 0 and 1 given by $c = \text{var}(x_t) \mid (\text{var}(x_t) + \text{var}(\delta_t))$, where $\delta_t = z_t - x_t$ is the exposure error. Again, a constant difference between the 2 exposure measures changes only the intercept.

Thus, the estimated regression coefficient is biased toward 0. In one pertinent case, $\beta_x = 0$, the naive estimate $\hat{\beta}_z$ is unbiased with $E(\hat{\beta}_z) = \beta_x = 0$; that is, under the classical error model, measurement error does not lead to spurious associations if there is truly no association. Random variation, of course, can produce such associations by chance, as it can absent measurement error. The probability of such false positive associations (the Type I error rate), however, remains the same.

For estimating effects of air pollution on mortality, realistic models have elements of both classical and Berkson error models. In general, the effect of such exposure errors is intermediate between the 2 extreme models. The effect of measurement error, therefore, is likely to depend on the direction and magnitude of the correlation of measurement errors with the measured exposures, not just on the variance of the measurement errors.

We now consider the more complex case of multipollutant models, which are often applied in an attempt to estimate the independent effect of one particular pollutant present in a mixture with other pollutants. For example, in an analysis of air pollution and mortality in Philadelphia, Kelsall and colleagues (1997) regress mortality on as many as 5 pollutants. Because little empirical evidence about the simultaneous errors in multiple pollutants is currently available, this section only lays a foundation that can inform the design of future studies, as discussed in the last section. Confining attention to the classical and the Berkson error cases, we again assume a linear regression model of the form given by equation (1), where x_t now represents a vector of exposure variables, with a corresponding vector of regression coefficients β_x; and z_t denotes a vector of measurements of each exposure variable. In the Berkson error case, the assumption that $E(x_t \mid z_t) = z_t$, for a vector of errors uncorrelated with z_t, still ensures that the estimates of the regression coefficients are unbiased, as in the univariate instance. But under the classical error model, the multiple regression extension is not so straightforward. As before, we assume that $E(z_t \mid x_t) = x_t$. To compute $E(x_t \mid z_t)$, let V denote the covariance matrix of x_t and let T denote the covariance matrix of $\delta_t = z_t - x_t$, and, as before, we assume that δ and x are independent. Then, the matrix generalization of the earlier result is that $\hat{\beta}_z = \hat{\beta}_x C$ where $C = T(T + V)^{-1}$. Now it is no longer true that $\beta_{zj} < \beta_{xj}$ for each component (j) and estimates of regression coefficients can be biased toward or away from the null; that is, positive associations can be produced even though the true coefficient for a particular component is 0 when the component is correlated with at least one component having a nonzero effect.

Table 1 illustrates the magnitude of bias that can result from regressing y_t on 2 predictors z_{1t} and z_{2t} instead of on x_{1t} and x_{2t}. This example might refer to estimating the effects of PM_{10} and O_3 on mortality when ambient values (z) instead of personal exposure (x) are available. We assume $z_{1t} = x_{1t}\delta_{1t}$ and $z_{2t} = x_{2t} + \delta_{2t}$, $V_{11} = \text{var}(x_{1t}) = V_{22} = \text{var}(x_{2t}) = 1$. The table presents the expected values for the estimated regression coefficients when the true values are both 1 ($\beta_{x1} = \beta_{x2} = 1$) for varying values of the correlation between x_{1t} and x_{2t}, the variances of δ_{1t} and δ_{2t}, and the correlation between the measurement errors δ_{1t} and δ_{2t}. At present, there is little empirical evidence about the nature or size of the correlations between pairs of pollutant measurements, and the table is intended to illustrate the consequences of measurement error in the 2-predictor model.

The first line of the table refers to an example in which there is no correlation between x_{1t} and x_{2t}; there is equal variability of the 2 exposure errors δ_{1t} and δ_{2t}; and these errors are not correlated. That is, the error in one predictor does not predict the error in the other. Here, there is an equal degree of attenuation in the coefficients for the

Table 1. Predicted Bias in Bivariate Regression Coefficients Under Different Covariance Structures for True Exposures and Measurement Errors When Both Variables Have True Effect: $\beta_{x1} = \beta_{x2} = 1.0$[a]

Row	Correlation (x_1, x_2)	Variance (δ_1)	Variance (δ_2)	Correlation (δ_1, δ_2)	$E(\hat{\beta}_{z1})$	$E(\hat{\beta}_{z2})$
1	0.0	1.0	1.0	0.0	0.50	0.50
2	0.5	1.0	1.0	0.0	0.60	0.60
3	−0.5	1.0	1.0	0.0	0.33	0.33
4	0.0	1.0	1.0	0.5	0.40	0.40
5	0.0	1.0	1.0	−0.5	0.67	0.67
6	0.0	0.5	2.0	0.0	0.67	0.33
7	0.5	0.5	2.0	0.0	0.71	0.53
8	0.5	0.5	2.0	0.3	0.66	0.27
9	0.5	0.5	2.0	0.5	0.64	0.21
10	0.5	0.5	2.0	0.7	0.64	0.14
11	0.5	0.5	2.0	−0.5	0.83	0.50
12	0.5	0.5	2.0	−0.7	0.91	0.57
13	0.5	0.5	2.0	−0.9	1.00	0.66

[a] We assume $\mathrm{Var}(x_1) = \mathrm{Var}(x_2) = 1$.

2 variables. With unequal variances but no correlation, that is, the sixth row, the degree of attenuation is greater for the variable with greater variance. If the exposures are correlated, but the errors are uncorrelated (the second and third rows), the 2 effect estimates are similarly altered with the direction of the effect depending on the sign of the correlation. Introducing correlation between the errors, that is, the fourth and fifth rows, has an effect that depends on the pattern of correlation. The bottom half of Table 1 shows more complex patterns with differing patterns of correlation and variation of the 2 errors. Some of the scenarios introduce substantially different effects of the 2 variables, but none yield effect estimates above the true value of 1, even with more extreme differences in error variances or the 2 correlations.

Table 2 also addresses the consequences of measurement error in a 2-variable model, but in this example only 1 variable (x_2) has a true effect; the other exposure x_1 has no effect on the health outcome y. Either correlation between x_{1t} and x_{2t} or their errors can introduce an apparent effect of x_1 on y. Some scenarios of variance and correlation even bring the apparent effects of the 2 variables quite close (eg, the 10th and 11th rows), but in every case, including more extreme situations than shown, the estimate for the true predictor (β_2) is always larger than for the null predictor (β_1).

Some general conclusions can be offered concerning multipollutant models under this simple, classical error model.

1. There is a general tendency for $\beta_{zj} < \beta_{xj}$ if all $\beta_{xj} > 0$.
2. The degree of attenuation of each coefficient depends, in large part, on its measurement error variance relative to the variance of the true exposure—that is, T_{jj}/V_{jj}. Thus, the coefficients for variables that are measured with considerable error will be attenuated more than those of variables with less error.
3. Depending on the correlation structure of the attenuation matrix C, some of the effect of one variable, β_{xj}, may be transferred to the estimate of another variable's effect, $\hat{\beta}_{zk}$. Such transfers of effect are generally from a more poorly measured variable to a better measured variable. For such transfers to be large, however, the true exposure variables or their measurement errors need to be substantially correlated.
4. As a consequence of conclusion (3), the estimate of a parameter can be biased away from the true value. However, this type of bias generally arises only with a very strong negative correlation between the measurement errors (eg, the 9th to 11th rows of Table 2).
5. Also as a consequence of (3), there will generally be spurious associations for a variable x_j that, in fact, has no effect only if x_j is substantially correlated with one or more variables that actually have an effect. Generally, the correlation among the errors has a larger influence on the bias than the correlation among the true pollutant levels.

Table 2. Predicted Bias in Bivariate Regression Coefficients Under Different Covariance Structures for True Exposures and Measurement Errors When Only One Variable Has True Effect: $\beta_{x1} = 0$, $\beta_{x2} = 1.0$[a]

Row	Correlation (x_1, x_2)	Variance (δ_1)	Variance (δ_2)	Correlation (δ_1, δ_2)	$E(\hat{\beta}_{z1})$	$E(\hat{\beta}_{z2})$
1	0.0	0.5	2.0	0.0	0.00	0.33
2	0.0	0.5	2.0	0.5	−0.12	0.35
3	0.0	0.5	2.0	−0.5	0.12	0.35
4	0.5	0.5	2.0	0.0	0.06	0.29
5	−0.05	0.5	2.0	0.0	−0.06	0.29
6	0.5	0.5	2.0	0.3	−0.01	0.28
7	0.5	0.5	2.0	0.5	−0.07	0.29
8	0.5	0.5	2.0	0.7	−0.15	0.29
9	0.5	0.5	2.0	−0.5	0.17	0.33
10	0.5	0.5	2.0	−0.7	0.21	0.36
11	0.5	0.5	2.0	−0.9	0.26	0.39

[a] We assume $\mathrm{Var}(x_1) = \mathrm{Var}(x_2) = 1$.

These conclusions are obtained from and therefore pertain to the classical linear regression model with 2 predictors, assuming z_t is a surrogate for x_t (nondifferential errors). The actual exposure measurement situation in the air pollution–mortality context is obviously more complex. First, log-linear, not linear, models are used, although the degree of nonlinearity is usually small in mortality studies. Second, the measurement errors are not purely of the classical, nondifferential type. For example, the degree of error for gaseous pollutants may depend on temperature or other covariates. Finally, errors may be multiplicative rather than additive. Nonetheless, the linear regression with classical measurement error is a leading case that provides insight into the major possible consequences of exposure errors.

FRAMEWORK FOR ASSESSING MEASUREMENT ERROR EFFECTS

Building on the fundamental concepts underlying statistical models of exposure measurement error, we focus here on the specific log-linear regressions used for assessing the pollutant-mortality association, controlling for weather variables. We identify 3 major components of measurement error and present a statistical framework for evaluating their potential effects on the estimated pollutant-mortality associations. The discussion below is based on the premise that the ideal investigation of the health effects of air pollution would be conducted at the individual level with measurements of personal exposure to pollutants. Exposure and mortality data, however, are generally available only after aggregation to a municipal level; little or no data from indoor air monitoring are available. Finally, air pollutant measurements are imprecise, and this imprecision has consequences for estimates of pollutant effects on mortality, as described in the previous section.

To investigate the effects of exposure error in the log-linear regressions widely used to assess the pollutant mortality association, consider the following model for an individual's risk of mortality:

$$\lambda_{it} = \lambda_{0it} \exp(x_{it}\beta_x) \tag{4}$$

where λ_{it} is the risk of death for person i on day t; λ_{0it} is that individual's baseline risk in the absence of exposure, that is, $x_{it} = 0$; and $\exp(x_{it}\beta_x)$ is the relative risk of death associated with the explanatory variables. Let $y_{it} = 1$ if person i dies on day t and 0 if he does not. We typically observe the total number of deaths for a population $y_t = \sum_{i=1}^{n_t} y_{it}$ where $n_t \approx n$ is the population size on day t. By (4), the expected total numbers of deaths λ_t in a community is

$$\lambda_t = Ey_t = \sum_i \lambda_{it} = \sum_i \lambda_{0it} \exp(x_{it}\beta_x). \tag{5}$$

In analyzing population-level data on mortality and air pollution, log-linear regressions of the form

$$\lambda_t = \exp(s(t) + z_t\beta_z + u_t\beta_u) \tag{6}$$

have been fit where $s(t)$ is an arbitrary but smooth function of time introduced to control for the confounding of longer-term trends and seasonality, z_t is the average of multiple monitor measurements of ambient pollution measurement

for day t, and u_t are other possible confounders such as temperature and dew point on the same and previous days.

If the regression coefficient β_x for a pollutant in the personal risk model (4) is the target for inference, how closely do estimates of β_z from model (6) approximate β_x? Below, we identify potential sources of bias in $\hat{\beta}_z$ as an estimate of β_x, using the concepts of Berkson and classical measurement error.

Figure 1 poses a model of the relationship between the personal exposure to a pollutant x_{it} for person i on day t and the available ambient values z_t measured with error by monitors. Assuming, for simplicity, a high degree of spatial homogeneity in ambient levels, personal exposure is contributed to by z_t^*, the true outdoor level; and by w_{it}, the indoor level, which is also influenced by z_t^* from penetration of the pollutant in outdoor air into indoor spaces. For example, personal exposure to PM_{10} is determined by the time spent outdoors, the concentration during that time, and by the concentrations in indoor environments that are determined by indoor sources such as cigarette smoking and the penetration of particles indoors, as air is exchanged between the outdoor and indoor environments. Figure 1 further shows that personal risk of dying is influenced by a person's baseline risk in addition to the unobserved personal exposure to pollutant x_{it}. Only the measured ambient pollution data, shown in a rectangular box, are actually observed.

In considering the consequences for $\hat{\beta}_z$, as an estimate of β_x, of having an imprecise measure of ambient pollution z_t, rather than actual personal exposure x_{it}, it is useful to

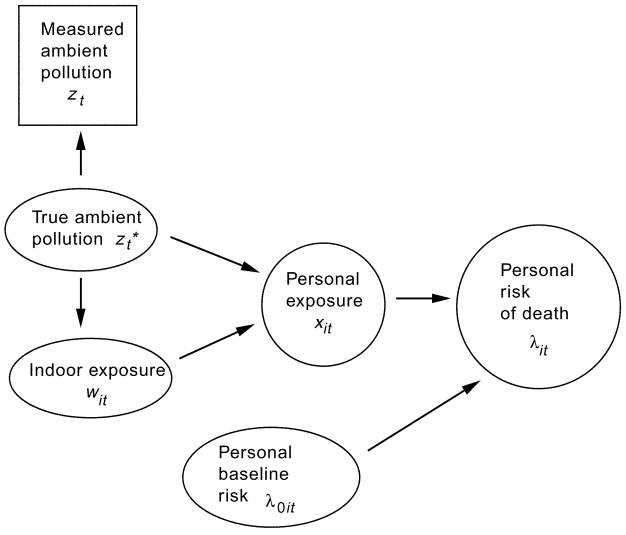

Figure 1. Ambient measured pollution level z_t as related to true ambient level z_t^*, indoor exposure w_{it}, personal exposure x_{it}, and risk of death λ_{it} assuming spatial homogeneity in ambient levels.

begin by decomposing the pollution measurement difference between x_{it} and z_t into 3 components:

$$x_{it} = z_t + (x_{it} - \bar{x}_t) + (\bar{x}_t - z_t^*) + (z_t^* - z_t). \quad (7)$$

Here, $(x_{it} - \bar{x}_t)$ is the error due to having aggregated rather than individual exposure data; $(\bar{x}_t - z_t^*)$ is the difference between the average personal exposure and the true ambient pollutant level; and $(z_t^* - z_t)$ represents the difference between the true and the measured ambient concentration.

The first term, $(x_{it} - \bar{x}_t)$, is an example of Berksonian error, so that in a simple linear model, having aggregate rather than individual exposure does not itself lead to bias into the regression coefficient. The second term, $(\bar{x}_t - z_t^*)$, is not Berksonian and is likely to be a source of bias. The final term, $(z_t^* - z_t)$, is largely of the Berkson type if the average of the available monitors z_t is an unbiased estimate of the true spatially averaged ambient level z_t^*.

We can now further study the effects of these 3 terms on risk estimation by substituting the decomposition in equation (7) into equation (5). After some straightforward calculations, the expected number of deaths on day t can be written

$$\mathrm{E}y_t = \exp\big(\log(n_t\bar{\lambda}_{0t}) + z_t\beta_x + [(\bar{x}_t^{(w)} - \bar{x}_t) + (\bar{x}_t - z_t^*) + (z_t^* - z_t)]\beta_x\big). \quad (8)$$

Here β_x is the personal log relative risk of interest from equation (5). Note the approximation of equation (8) retains only linear terms in the expansion of an exponential function. The second-order terms are an order of magnitude smaller and are ignored to simplify the exposition. For studies of particulate pollution effects on mortality, the effect sizes are on the order of only a percent or two, so that ignoring second-order terms should not qualitatively change the results. In studies of morbidity, higher-order terms may be more important.

The total baseline risk, $(n_t\bar{\lambda}_{0t})$, almost certainly varies smoothly over time, since it is an average risk over a large population. Hence, it will be appropriately controlled for in log-linear regressions by inclusion of the smooth $s(t)$ in equation (6). We now consider $z_t\beta_x$ and the 3 components of error in turn.

The first error term, $(\bar{x}_t^{(w)} - \bar{x}_t)$, is proportional to the difference between the baseline-risk-weighted average personal exposure and the unweighted average personal exposure, where risk refers to the probability of mortality during the reference period for exposure. It derives from the Berkson error $(x_{it} - \bar{x}_t)$ and produces no bias in the linear, unaggregated model. This difference, which is due

to risk weighting in our log-linear model with person-specific baseline risks, is likely to be small and to vary slowly over time. Hence, it can be adequately controlled by inclusion of the smooth function $s(t)$ in the log-linear regression of y_t on z_t. One scenario in which this difference would vary from day to day and therefore not be adequately controlled would occur if the more frail individuals were to follow pollution reports (or a correlate such as weather) and reduce their exposures to ambient air on high pollution days by, for example, staying indoors. Current warning systems for air pollution alerts are intended, in fact, to reduce exposures of susceptible persons in this fashion.

The second error term, $(\bar{x}_t - z_t^*)$, is non-Berksonian and has the greatest potential to introduce bias in the estimate $\hat{\beta}_z$ when z_t^* is correlated with $(\bar{x}_t - z_t^*)$. Even if the terms are uncorrelated so that $\hat{\beta}_z$ will be a roughly unbiased estimate of β_x, it will reduce efficiency relative to a study in which x_t is available, since z_t and $(\bar{x}_t - z_t)$ share the same coefficient in equation (8).

The difference, $(\bar{x}_t - z_t^*)$, between average personal exposures and the true ambient value can be analyzed further by considering an individual personal exposure x_{it}. Because the exposure of individual i on day t derives either from indoor or ambient sources, we can write $x_{it} = \alpha_{it} z_t^* + (1 - \alpha_{it}) I_{it}$ where I_{it} is the concentration of pollutant generated by indoor sources such as tobacco smoke and pets, and α_{it} is the individual's fraction of exposure from ambient sources that takes place either outdoors or results from penetration of ambient pollution indoors. It follows that $\bar{x}_t = \bar{\alpha}_t z_t^* + \bar{I}_t$ where $\bar{I}_t = \sum_i (1 - \alpha_{it}) I_{it}/n_t$. That is, the average personal exposure is proportional to the ambient level offset by the effects of the population average of the nonambient indoor sources.

Wilson and Suh (1997) have argued that the daily population average concentrations of fine particles derived from indoor sources $\bar{I}_t$ are approximately independent of ambient levels z_t across time. When this is true, failure to measure indoor sources will not introduce further bias in the estimation of β_x because the deviations due to indoor air exposure are a second example of Berkson error, and these errors will tend to cancel one another out when averaged over the population. Nevertheless, z_t^* is only proportional to $\bar{x}_t$, so that, even if $\bar{\alpha}_t$ varied little over time ($\alpha_t \approx \alpha$), the coefficient $\hat{\beta}_z$ from a regression of y_t on z_t^* would estimate $\alpha\beta_x$, not β_x. Hence, if 20% of daily exposure results from indoor sources independent of ambient levels, the regression on ambient levels will yield coefficients that are roughly 20% smaller than would have occurred with actual personal exposures. This may be the appropriate coefficient for policy makers seeking an estimate of the effect of an inarguable measure of ambient levels. This assumes, however, that particles from indoor sources and outdoor sources are identical; that is, they are similar in composition and toxicity. If this is not the case, then the 2 types of particles would be more appropriately treated as separate pollutants, and the personal exposure measure desired would be $\alpha_{it} z_t^*$, the personal exposure to particles from outdoor sources. Studies using sulfates as a tracer for particles from outdoor sources indicate that indoor/outdoor ratios are less than 1. Since people spend most of their time indoors, this suggests that α_{it} will be less than 1, and that the second term in equation (8) will be negatively correlated with z_t and will bias the estimated coefficient downward. This also illustrates that the model is not restricted to cases where $E(x) = E(z)$.

The last of the 3 error terms in equation (8), $(z_t^* - z_t)$, represents the instrument measurement error in the ambient levels; like $(x_{it} - \bar{x}_t)$, it is close to the Berkson type. This term would tend to be cancelled out by spatial averaging across multiple, unbiased ambient monitors in a region. For example, Kelsall and colleagues (1997) averaged daily total suspended particles data from up to 9 monitors in their analysis of effects of particles on mortality in Philadelphia. However, in many cities there are only one or a few monitors operating at a time. Even with a small number of monitors, longer-term drift in instruments will not substantially affect estimates of β_x because the time-series models control for such trends by including $s(t)$ in equation (6). For this final error term to cause substantial bias in $\hat{\beta}_z$, the error, $(z_t^* - z_t)$, must be strongly correlated with z_t at shorter time scales. Further investigations of this correlation in cities with many monitors are warranted.

To summarize, we have discussed 3 components of measurement error: (1) an individual's deviation from the risk-weighted average personal exposure; (2) the difference between the average personal exposure and the true ambient level; and (3) the difference between the measured and the true ambient levels, which includes spatial variation and instrument error. Our analysis argues that the first and third components are of the Berkson type and, therefore, are likely to have smaller effects on the relative risk estimates. However, the second component can be a source of substantial bias, if, for example, there are short-term associations of the contributions of indoor sources with ambient concentrations. The following simple analysis of the PTEAM data illustrates how we can further study the effects of the most important second component.

EVALUATING POTENTIAL MEASUREMENT ERROR BIAS IN RELATIVE RISK ESTIMATES

The framework in combination with data on the components of error can be used to evaluate and possibly adjust for the consequences of exposure measurement error. We use one of the few available data sets with ambient and personal measurements to illustrate one approach. We begin by using daily measurements of personal exposure for 178 persons followed in the PTEAM study (Ozkaynak et al 1996) to quantify the difference between concentration measured by an ambient monitor and the average of personal exposures. We then present one approach for estimating the degree of bias in estimated PM_{10}-mortality regression coefficients $\hat{\beta}_z$. This coefficient is an estimate of the true relative risk for personal exposure β_x with data from 1 or a few ambient monitors rather than personal exposure data for PM_{10}.

PTEAM STUDY DATA

The PTEAM study (Mendelsohn and Orcutt 1979; Ozkaynak et al 1996) generated a daily measurement of personal exposure to PM_{10} for a sample of 178 nonsmoking residents of Riverside, California, aged 10 years or older, for the period September 22 through November 9, 1990. In addition, a daily average PM_{10} value from an ambient monitor positioned near the homes was also collected; Pellizzari and Spengler (1990) provide details on the methods used to collect these data.

We use the PTEAM study data to estimate the correlation between the daily PM_{10} concentration for the ambient monitor z_t and the difference between the average personal exposure and concentration measured by the ambient monitor, $(\bar{x}_t - z_t)$. These estimates correctly account for the varying number of observations on a given day. Note that equation (8) includes a weighted average of personal exposures, with weights determined by the baseline personal risk for each individual. In the PTEAM study, those weights are unavailable; hence, an unweighted average is used. Figure 2 displays a time-series plot of the daily ambient values and the average personal exposures. The correlation across time of these 2 series is estimated to be 0.58 (95% confidence interval [CI], 0.35 to 0.74). This is much greater than the more widely cited cross-sectional correlation from this study. The corresponding correlation across time between the ambient monitor concentrations and the daily differences between the personal and ambient values is –0.63 with 95% CI, –0.77 to –0.42. Hence, the hypothesis (the measurement error, $[\bar{x}_t - z_t]$, is uncorrelated with z_t) is not consistent with the PTEAM study data. Some bias in the regression coefficient is therefore expected. Because the correlation of $(\bar{x}_t - z_t)$ and z_t is negative, the coefficient $\hat{\beta}_z$ in the regression on z_t will tend to underestimate the coefficient in the regression on $\bar{x}_t$ in a single-pollutant analysis.

ADDRESSING BIAS IN PM_{10}-MORTALITY REGRESSION COEFFICIENTS

The PTEAM study results or other, perhaps more appropriate, data sets on the difference between average risk-weighted personal exposure and ambient monitor concentrations can be used to estimate bias in the results of log-linear regression models.

If they had been available, we would have used the average personal exposure series, $\bar{x}_t$, for at-risk residents of each city in the standard log-linear regression model rather than z_t, as was used in the original analyses. We would then have compared the regression coefficients obtained when $\bar{x}_t$ is the predictor with those using z_t to assess the bias.

Obviously, $\bar{x}_t$ is not available except in special circumstances. From the PTEAM study data, however, shown in Figure 2 or similar data, we can estimate the relationship of $\bar{x}_t$ and z_t, for example, by assuming:

$$\bar{x}_t = \theta_0 + \theta_1 z_t + \varepsilon_t, \tag{9}$$

where θ_0 and θ_1 are the intercept and slope to be estimated from the available data. We can then use the fitted equation (9) to predict the unobserved $\bar{x}_t$ from the available z_t, and then use the predicted value $\hat{\bar{x}}$ as the desired exposure values when estimating the pollution-mortality relative risk β_x. The estimate of β_x has the simple form $\hat{\beta}_x = \hat{\beta}_z / \hat{\theta}_1$. This is one well-known approach to adjust for exposure measurement error called *regression calibration* (Carroll et al 1995). As an illustration, we have applied this strategy to a regression of daily mortality on ambient concentrations of PM_{10} for Riverside, California, for the period 1987 to 1994. We estimate $\hat{\theta}_0 = 59.95$ (se = 7.21), $\hat{\theta}_1 = 0.60$ (se = 0.080), and var(ε) = 22.4.

Calibration is easy to implement and apply. Its limitations are that CIs for $\hat{\beta}_x$ depend on large sample theory, and calibration does not extend easily to situations where multiple sources of information about the $\bar{x}_t$, z_t relationship are available.

It is simple, however, to overcome these possible limitations of calibration by using a simulated value $\bar{x}^*_t$ rather than the predicted value $\hat{\bar{x}}$ from equation (9). That is, we use equation (9) to simulate the average personal exposure, $\bar{x}^*_t$, from the ambient exposure, z_t, for a city or period of interest when $\bar{x}_t$ is not available, under the assumption that the estimated θ and var(ε) are applicable. This simulated series $\bar{x}^*_t$ is then used instead of z_t in the log-linear

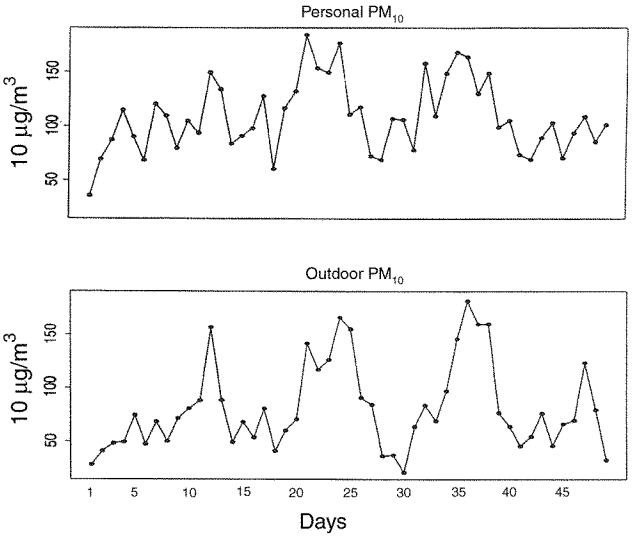

Figure 2. PTEAM Study. Daily time series of personal and central site concentrations of PM_{10} for Riverside, California, from September 22 to November 9, 1990 (Ozykaynak et al 1996).

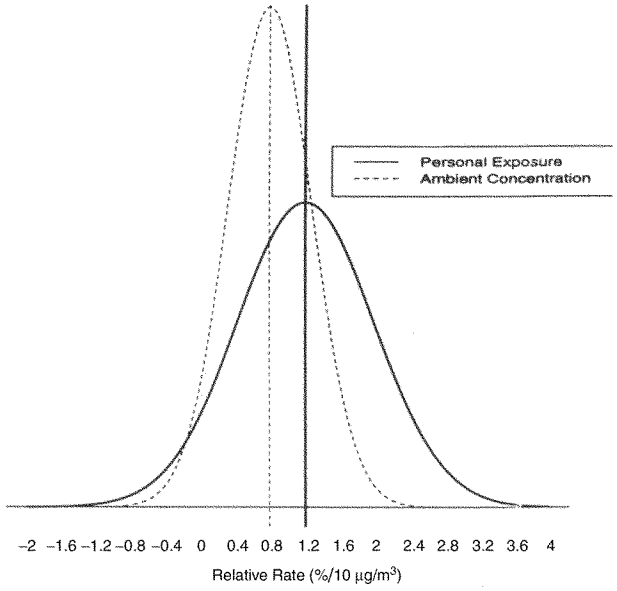

Figure 3. Distribution of personal exposure effects. The solid line is the distribution of the relative rate $\hat{\beta}_x$ obtained when the simulated series $\bar{x}_t$ of the total personal exposure is the predictor in the log-linear regression. The dotted line is a normal approximation of the distribution of the relative rate $\hat{\beta}_z$ obtained when the ambient concentration z_t in Riverside, California, is the predictor in the log-linear regression.

regression. The result is one estimate of β_x—call it $\hat{\beta}_x$. If we then repeatedly simulate $\bar{x}_t^*$ and fit the log-linear regression for each to obtain $\hat{\beta}_x$, we obtain a distribution of $\hat{\beta}_x$. The difference between the mean of the simulated $\hat{\beta}_x$ and the $\hat{\beta}_z$ derived from the log-linear regression of mortality on $\bar{z}_t^*$ is a measure of the bias resulting from having z_t rather than the average personal exposure for that city. By simulating $\bar{x}_t^*$ rather than using a fixed predicted value $\hat{\bar{x}}_t$, we properly account for nonlinearities and sources of variation in $\hat{\beta}_x$ and can extend the analysis to more complicated situations.

Figure 3 shows the distribution of the $\hat{\beta}_x$ for Riverside (solid curve). Also shown is the normal approximation of the likelihood function for the coefficient $\hat{\beta}_z$ from the log-linear regression of mortality directly on z_t (dotted curve). Solid and dotted lines are at the centers of these distributions. We find that the $\hat{\beta}_x$ have a mean 1.42% increase in mortality (95% interval: –0.11, 2.95) per 10-unit change in PM_{10}. In comparison, the estimate of β_z from the usual log-linear model (seen as a dashed vertical line in Figure 3) is $\hat{\beta}_z$ = 0.84% (95% interval: –0.06, 1.76). Hence, measurement error has biased the result toward the null. Second, the distribution of the $\hat{\beta}_x$ is more dispersed than the distribution of $\hat{\beta}_z$. This is because we have taken into account the variability due to having z_t, not $\bar{x}_t$. These results are very similar to the ones we obtain from calibration.

The calculation in the previous paragraph assumes the estimated relationship between $\bar{x}_t$ and z_t for the PTEAM study is the true one, and hence, we ignore a second component of uncertainty that occurs from estimating the relationship between $\bar{x}_t$ and z_t from the finite sample size of the PTEAM study data taken at 1 site and a particular time period. That is, even if we assume that the relationship between $\bar{x}_t$ and z_t is known, estimating the association of mortality with $\bar{x}_t$ is less precise than estimating it with z_t, given only z_t in that particular city. Of course, the relationship of $\bar{x}_t$ and z_t is not precisely known and needs to be quantified further. Dominici and colleagues in section 2 of this report provide a more complete analysis of the bias in $\hat{\beta}_z$ as an estimate of β_x using the PTEAM study and 4 other data sets and a more complete statistical model than described here. Their findings are qualitatively similar to those presented here. Finally, it is important to note that our assessment of bias assumes that the health effects of personal exposure to particles originating outdoors and indoors are the same. To assume otherwise would require substantially more detailed data and modeling.

SUMMARY AND RESEARCH RECOMMENDATIONS

The differences between true personal exposure for every individual (x_{it}) and measured ambient concentrations, averaged over a few fixed, imprecise monitors (z_t), are inherently complex, as is the effect of this exposure measurement error on estimates of pollution-mortality relative risks. Nonetheless, it is useful and imperative to analyze these effects in light of our current understanding of the measurement process. This section presents one framework for doing so. We distinguish 2 extremes of a continuum of types of measurement errors: Berkson and classical errors. The former is likely to create little bias in mortality-relative risk estimates; the latter has more serious consequences.

We posit a relative risk model in which an individual's hazard of death on a given day is expressed as a function of his or her personal exposure, which is decomposed to highlight 3 types of exposure errors. This model is then aggregated to produce the model for the expected total deaths in a population used in most time-series analyses. This model shows that a risk-weighted average personal exposure measure is the desired one, and we discuss the consequences of the widely used feasible alternative, ambient concentration. In contrast, differences between individual exposures on a given day and the risk-weighted average of personal exposures are an example of Berkson error and are not likely to cause substantial bias in coefficients from time-series mortality studies. Our analysis suggests that the largest biases in inferences about the mortality–personal exposure relative risk will occur due to the more complex errors between ambient and average personal exposure measures. If indoor sources produce particles of similar composition and toxicity as outdoor source particles, indoor sources may be a major component of this error. Finally, as an illustration we have used the best available data, that from the PTEAM study in Riverside, California, with both personal exposure and ambient time series to quantify the size of this error. Our analysis indicates that the coefficient obtained from regressing mortality on measured ambient level z_t is smaller than what we expect from regressing mortality on average personal exposure $\bar{x}_t$.

For tractability and clarity, we have conducted a first-order analysis of exposure errors and have ignored possible second-order and higher-order effects in which daily fluctuations in the variance of personal exposures across a population or in the covariations among the measurement errors could introduce additional biases. Second-order terms will be insignificant in studies of particulate effects on mortality where the first-order terms are on the order of percent. Such higher-order analyses for other studies of, for example, morbidity, are beyond the scope of this report and will require substantially more detailed models and data. It is, however, possible that higher-order effects are important; thus further investigation is necessary.

Epidemiologic research is necessarily limited by the quality of the health outcome and risk factor measurements (Vedal 1997). Time-series studies of the acute effects of air quality on mortality are subject to the limitations posed by the available measurements of pollution levels. The generic criticism—that measurement errors render the results of such time-series models uninterpretable—is incorrect. The consequences of measurement error can be quantified, although only a few informative data sets are presently available. Further differences between the average personal exposure and ambient measurements are the most likely source of substantial bias. We suggest that data should be collected for comparison of risk-weighted average personal exposure with ambient levels in several cities with varying degrees of spatial heterogeneity in ambient levels, population composition, and indoor pollution sources. Given such data, models such as those summarized in this section and by Dominici and colleagues in section 2 of this report can be used to quantify more precisely the biases due to pollutant measurement errors.

Section 2 focuses on the effects on relative risk estimates of using z_t, measured ambient particle levels, rather than x_{it}, and actual personal exposures in log-linear regressions. Such effects are important from a scientific perspective to quantify the health risks of exposure to particulate pollution. From a regulatory perspective, the effect of having the imprecise z_t rather than the true ambient value z_t^* may be of greater interest, since what may or may not be regulated further are ambient levels. A more detailed error analysis of the $z_t - z_t^*$ difference would investigate the spatial variation in particulate levels and the way that the number of monitors used to calculate z_t reduced this source of measurement error.

The analyses of sections 3 and 4 focus on measurement error in a single pollutant measure, PM_{10}. Simultaneous errors in several pollutants can complicate the analysis. As clearly demonstrated, however, qualitative biases—that is, changes in the sign of a coefficient—can occur only when the measurement errors for different pollutants are highly correlated with one another. This level of correlation might arise if 2 or more pollutants are measured by the same instrument (eg, different fractions of PM) or if multiple instruments are housed in the same location, which is subject to atypical exposure patterns. The possibility of this level of correlation nevertheless requires detailed

investigation because in this case the findings of epidemiologic studies could be misleading. Personal exposure studies that collect multiple exposures can provide the necessary data to investigate the effects of co-occurring errors using straightforward extensions of the approaches outlined in sections 3 and 4.

The measurement error framework and the illustrative calculations in section 2 make apparent several open questions and opportunities for additional data collection that would enable more accurate quantification of the effects of measurement error in assessing the air pollution–mortality relationship. In relation to single-pollutant models, we consider that the 2 most important questions are:

- Is the average personal exposure to pollutants from indoor sources correlated with ambient levels over time?
- Does the difference between baseline risk-weighted average exposure and population average exposure vary slowly over time?

For models with multiple pollutants, the additional key question is:

- How do the components of error identified in equation (5) co-vary across pollutants? For example, how do the differences between actual ambient levels and the measured levels correlate across the different pollutants, and how do these differences depend on the true values of other pollutants or covariates?

Wilson and Suh (1997) have conducted a meta-analysis of data from multiple sites and conclude, in answer to the first question above, that concentrations of fine particles originating from indoor sources are independent of ambient levels over time. To confirm this finding and to address the remaining key questions, additional research is warranted. It would be highly informative if, in several cities with diverse pollution sources and patterns, a stratified sample of the population were drawn with one stratum representing the entire population and the second representing the frail subgroup. Daily measurements of personal exposure and indicators of indoor sources would be collected for multiple pollutants for each person. Ambient levels would also be monitored. Decisions about the numbers of persons within each subgroup and the numbers of days of monitoring for each person would be made based on preliminary analyses of data from 1 city.

REFERENCES

Armstrong BK, Saracci R, White E, eds. 1992. Principles of Exposure Measurement in Epidemiology. Oxford University Press, New York NY.

Carroll RJ, Ruppert D, Stefanski LA, eds. 1995. Measurement Error in Nonlinear Models. Chapman & Hall, London, England.

Carroll RJ, Spiegelman CH, Lan KKG, Bailey KT, Abbott RD. 1984. On errors in variables for binary regression models. Biometrika 71:19–25.

Dockery DW, Pope CA III. 1994. Acute respiratory effects of particulate air pollution. Annu Rev Public Health 15:107–132.

Dockery DW, Pope CA III, Xu X, Spengler JD, Ware JH, Fay ME, Ferris BG Jr, Speizer FE. 1993. An association between air pollution and mortality in six US cities. N Engl J Med 329:1753–1759.

Dominici F, Zeger S, Samet J. 1999. A Measurement Error Correction Model for Time-Series Studies of Air Pollution and Mortality. Technical Report. Johns Hopkins University, Baltimore MD.

Ferris BG Jr, Speizer FE, Spengler JD, Dockery DW, Bishop YM, Wolfson M, Humble C. 1979. Effects of sulfur oxides and respirable particles on human health: Methodology and demography of populations in study. Am Rev Respir Dis 120:767–779.

Fuller WA, ed. 1987. Measurement Error Models. John Wiley & Sons, New York NY.

Goldstein IF, Landovitz L. 1977a. Analysis of air pollution patterns in New York City–I. Can one station represent the large metropolitan area? Atmos Environ 11:47–52.

Goldstein IF, Landovitz L. 1977b. Analysis of air pollution patterns in New York City–II. Can one aerometric station represent the area surrounding it? Atmos Environ 11:53–57.

Janssen NA, Hoek G, Brunekreef B, Harssema H, Mensink I, Zuidhof A. 1998. Personal sampling of particles in adults: Relation among personal, indoor, and outdoor air concentrations. Am J Epidemiol 147:537–547.

Kelsall JE, Samet JM, Zeger SL, Xu J. 1997. Air pollution and mortality in Philadelphia, 1974–1988. Am J Epidemiol 146:750–762.

Lipfert FW. 1997a. Air pollution and human health: Perspectives for the '90s and beyond. Risk Anal 17:137–146.

Lipfert FW. 1997b. Clean air skepticism. Science 278:19–20.

Lipfert FW, Wyzga RE. 1997. Air pollution and mortality: The implications of uncertainties in regression modeling and exposure measurement. J Air Waste Manage Assoc 47:517–523.

Mendelsohn R, Orcutt G. 1979. An empirical analysis of air pollution dose-response curves. J Environ Econ Manage 6:85–106.

National Research Council. 1985. Commission on Life Sciences, Board on Toxicology and Environmental Health Hazards, Committee on the Epidemiology of Air Pollutants. Epidemiology and Air Pollution. National Academy Press, Washington DC.

National Research Council. 1991. Committee on Advances in Assessing Human Exposure to Airborne Pollutants. Human Exposure Assessment for Airborne Pollutants: Advances and Opportunities. National Academy Press, Washington DC.

Navidi W, Thomas D, Stram D, Peters J. 1994. Design and analysis of multilevel analytic studies with applications to a study of air pollution. Environ Health Perspect (Suppl 8) 102:25–32.

Ozkaynak H, Xue J, Spengler J, Wallace L, Pellizzari E, Jenkins P. 1996. Personal exposure to airborne particles and metals: results from the Particle TEAM Study in Riverside, California. J Expo Anal Environ Epidemiol 6:57–78.

Pellizzari E, Spengler J. 1990. Particle Total Exposure Assessment Methodology (PTEAM): Pilot Study, Vol II: Protocols for Environmental Sampling and Analysis. Work Plan for EPA Contract No 68-02-4544, EPA Work Assignment 67, CARB Agreement No A833-060. US Environmental Protection Agency, Office of Air Quality Planning and Standards, Research Triangle Park NC.

Pierce DA, Stram DO, Vaeth M. 1990. Allowing for random errors in radiation dose estimates for the atomic bomb survivor data. Radiat Res 123:275–284.

Schwartz J, Dockery DW, Neas LM. 1996. Is daily mortality associated specifically with fine particles? J Air Waste Manage Assoc 46:927–939.

Shy CM, Kleinbaum DG, Morgenstern H. 1978. The effect of misclassification of exposure status in epidemiological studies of air pollution health effects. Bull NY Acad Med 54:1155–1165.

Snedecor GW, Cochran WG, eds. 1980. Statistical Methods. The Iowa State University Press, Ames IA.

Spiegelman D, McDermott A, Rossner B. 1997. Regression calibration method for correcting measurement-error bias in nutritional epidemiology. Am J Clin Nutr (Suppl 4) 65:1179S–1186S.

Thomas D, Stram D, Dwyer J. 1993. Exposure measurement error: Influence on exposure-disease relationships and methods of correction. Annu Rev Public Health 14:69–93.

Vedal S. 1997. Ambient particles and health: Lines that divide. J Air Waste Manage Assoc 47:551–581.

Wallace L. 1996. Indoor particles: A review. J Air Waste Manage Assoc 46:98–126.

Willett W. 1998. Correction for the effects of measurement error. In: Nutritional Epidemiology. Oxford University Press, New York NY.

Wilson WE, Suh HH. 1997. Fine particles and coarse particles: Concentration relationships relevant to epidemiologic studies. J Air Waste Manage Assoc 47:1238–1249.

Section 2: A Measurement Error Model for Time-Series Studies of Air Pollution and Mortality

Francesca Dominici, Scott L Zeger, and Jonathan M Samet

ABSTRACT

One barrier to interpreting the observational evidence concerning the adverse health effects of air pollution for public policy purposes is the measurement error inherent in estimates of exposure based on ambient pollutant monitors. Exposure assessment studies have shown that data from monitors at central sites may not adequately represent personal exposure. Thus, the exposure error resulting from using centrally measured data as a surrogate for personal exposure can potentially lead to bias in estimates of the health effects of air pollution.

This section of the Investigators' Report presents a multistage Poisson regression model for evaluating the effects of exposure measurement error on estimates of effects of ambient particulate matter (PM)* on mortality in time-series studies. To implement the model, we have used 5 validation data sets on personal exposure to PM less than 10 μm in aerodynamic diameter (PM_{10}). Our goal is to combine data on the associations between ambient concentrations of PM and mortality for a specific location, with the validation data on the association between ambient and personal concentrations of PM at the locations where data have been collected. We use these data in a model to estimate the relative risk of mortality associated with estimated personal exposure concentrations and compare this estimate with the risk of mortality estimated with measurements of ambient concentration alone. We apply this method to data comprising daily mortality counts, ambient concentrations of PM_{10} measured at a central site, and temperature for Baltimore, Maryland, from 1987 to 1994. We have selected our home city of Baltimore to illustrate the method; the measurement error correction model is general and can be applied to other appropriate locations.

Our approach uses a combination of (1) a generalized additive model with log link and Poisson error for the mortality–personal exposure association, (2) a multistage linear model to estimate the variability across the 5 validation data sets in the personal–ambient exposure association, and (3) data augmentation to address the uncertainty resulting from the missing personal exposure time series in Baltimore. In the Poisson regression model, we account for smooth seasonal and annual trends in mortality using smoothing splines. Taking into account the heterogeneity across locations in the personal–ambient exposure relationship, we quantify the degree to which the exposure measurement error biases the results toward the null hypothesis of no effect, and estimate the loss of precision in the estimated health effects due to indirectly estimating personal exposures from ambient measurements.

INTRODUCTION

Pollution of outdoor air is a public health concern throughout the world. In the last decade, many epidemiologic studies have shown an association between measurements of ambient concentrations of PM_{10} and nonaccidental daily mortality counts (Dockery and Pope 1994; Schwartz 1995; Bascom et al 1996a,b; Dominici et al 2000). These studies suggest that daily rates of morbidity and mortality from respiratory and cardiovascular diseases increase with levels of particulate air pollution even at levels well below the current National Ambient Air Quality Standard (NAAQS) for PM in the United States.

One scientific objective of risk assessment of particulate air pollution is an estimation of the increase in risk of mortality per unit increase in personal exposure to particulates. Epidemiologic studies, however, rarely obtain personal exposures and instead use measurements of ambient concentrations obtained typically from one or a few monitors stationed in the region where aggregate rates of morbidity or mortality are assessed. When exposures are

* A list of abbreviations and other terms appears on page 13.

The National Morbidity, Mortality and Air Pollution Study: Methods and Methodologic Issues, Part I of Health Effects Institute Research Report 94, includes an Investigators' Report, a Preface, a Commentary by the Health Review Committee, and an HEI Statement about the research project. Correspondence concerning this section may be addressed to Dr Francesca Dominici, Department of Biostatistics, Johns Hopkins School of Public Health, 615 North Wolfe Street, Baltimore MD 21205-2179.

Although this document was produced with partial funding by the United States Environmental Protection Agency under Assistance Award R824835 to the Health Effects Institute, it has not been subjected to the Agency's peer and administrative review and therefore may not necessarily reflect the views of the Agency, and no official endorsement by it should be inferred. The contents of this document also have not been reviewed by private party institutions, including those that support the Health Effects Institute; therefore, it may not reflect the views or policies of these parties, and no endorsement by them should be inferred.

measured with error, the power of such epidemiologic analyses is reduced (Carroll et al 1995).

In a recent nationwide study of the ways that people spend their time, based on interviews with 9,386 respondents in 1993 to 1994 (Robinson and Nelson 1995), US residents were found to spend 87.2% of time indoors, 7.2% in or near vehicles, and only 5.6% outdoors. Consequently, exposure to air pollution takes place in multiple environments over time and in multiple locations, and accurately estimating the actual personal exposures responsible for disease or death can be daunting. Exposure assessment studies (Lioy et al 1990; Mage and Buckley 1995; Ozkaynak et al 1996; Wallace 1996; Janssen et al 1997, 1998) show that data from monitors at central sites may be only weakly correlated with personal exposures, indicating that using centrally measured data as a surrogate can lead to bias in estimates of the health effects of air pollution.

We have developed a model for investigating differences in relative risk estimates that arise from using ambient concentration measurements rather than personal exposure measurements. Let X_t represent average personal exposure to particulate pollution, Z_t be the measured ambient concentration of particulates, and Y_t be the numbers of deaths in a region of interest (eg, Baltimore, Maryland). We suppose that the regression of interest has the log-linear form $\mathrm{E}[Y_t] = \exp(X_t \beta_x + \textit{confounders})$ so that β_x is the log relative risk of death associated with a unit change in average personal exposure. Suppose, however, that only time-series data from monitors Z_t rather than X_t are available for Baltimore. Suppose further that, for several other locations and/or time periods, we have limited measurements of both X_t and Z_t, so that a model can be constructed for the relationship of average personal exposures and ambient concentrations, taking account of variations within and across locations.

This model combines the log-linear model for Y_t given X_t with a measurement error model for X_t given Z_t to make inferences about β_x. The parameter β_x is important from a scientific and etiologic perspective since it quantifies the human risk of actual particulate pollution. As discussed below, the direct regression of Y_t on Z_t giving β_z is also of interest from a regulatory perspective, because only ambient concentrations are currently regulated.

Our modeling approach incorporates relevant data from 5 separate epidemiologic studies and properly accounts for heterogeneity in the $X_t - Z_t$ relationship across studies. This hierarchical extension also allows us to apply the measurement model to a new site such as Baltimore and to include uncertainty about the $X_t - Z_t$ relationship in the estimation of β_x for that new site.

More specifically, at the first stage of the measurement model we use a Poisson regression model to describe the association between daily mortality and the population-average personal exposure in Baltimore, which is a missing predictor. At the second stage, we use the supplemental information about personal exposure from the available exposure studies to model the association between average personal exposure and ambient concentrations. We use a combination of Bayesian hierarchical modeling (Lindley and Smith 1972; Morris and Normand 1992) and data augmentation (Tanner 1991) to estimate β_x, the log relative rate of mortality associated with average personal exposures to PM_{10} for Baltimore.

A hierarchical model is a flexible tool for modeling variability across studies of the relationship of personal and ambient exposure concentrations. Hierarchical multivariate regression models with missing predictors for both continuous and categorical data have been developed by Dominici (2000) and Dominici and colleagues (1997, 1999). Data augmentation can be used to account for uncertainty appropriately in estimates of the log relative rate of mortality resulting from the missing personal exposure data. Computationally, data augmentation can be handled conveniently using Markov chain Monte Carlo (MCMC) techniques (Tanner and Wong 1987; Gelfand and Smith 1990; Tanner 1991; Spiegelhalter et al 1994), which we adopt here.

Regression calibration models have previously been used to account for measurement error in nonlinear models (Carroll et al 1995). Armstrong (1985) suggested regression calibration for linear models, and Rosner, Willett, and Spiegelman (1989) developed regression calibration methods for logistic regression. Carroll and Stefanski (1990) applied regression calibration to generalized additive models with measurement error, using quasi-likelihood methods to handle predictors measured with error. Prentice (1982) and Clayton (1991) discuss the covariate measurement error problem for the proportional hazards model. See Fuller (1987) and Stefanski (1985) for a detailed discussion of measurement error in linear and nonlinear models, respectively. Dellaportas and Stephens (1997), Richardson and Gilks (1993), and Mallick and Gelfand (1994) have developed Bayesian measurement error approaches.

The Data Sources section describes the database for Baltimore, Maryland, and for the 5 studies reporting personal concentrations and ambient measurements of PM_{10}. These represent the best data on personal exposures published in the epidemiologic literature available to us. In the Methods section, we describe a generalized additive model with log link and Poisson error that we used to estimate β_z—the log

relative rate of mortality associated with a unit change in measured ambient concentrations of PM_{10}—directly. The semiparametric components are smoothing splines used to describe long-term fluctuations in mortality due to seasonality, changing population health, and confounding effects of weather. We first used the model to estimate β_z for Baltimore for the period 1987 to 1994.

The General Framework subsection for handling the problem of measurement error includes a description of the main components of error. Our simple hierarchical model for X_t given Z_t has been incorporated with the log-linear model to obtain a procedure for estimating β_x. This procedure takes into account the uncertainty arising because the personal exposures were not observed directly and ambient levels were used as a surrogate. The results of our analysis include a comparison with a non-Bayesian alternative based on a 2-stage regression calibration approach and an assessment of sensitivity to the prior distribution and modeling assumptions.

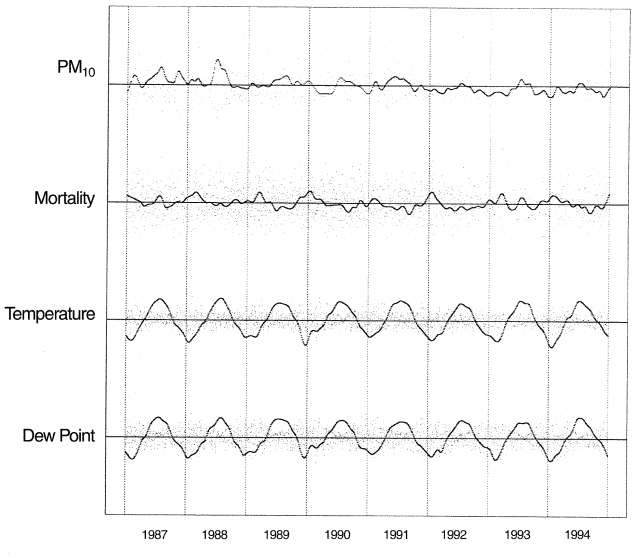

Figure 1. PM_{10} (μg/m^3), mortality, temperature and dew point for daily time series on Baltimore for the 8-year period 1987–1994. Solid lines are smoothing splines with 6 *df*/year; points are residuals of the raw data with respect to the smoothing splines.

DATA SOURCES

The Baltimore data include daily mortality, temperature, dew point, and particulate pollution concentrations for the 8-year period 1987 to 1994 (Figure 1). In addition to the data for Baltimore, we have assembled data from 5 studies of personal exposure and ambient concentrations of PM_{10} (Table 1). The 5 studies had heterogeneous goals, population characteristics, sampling schemes, observation periods, and locations as follows:

1. In the Particle Total Exposure Assessment Methodology (PTEAM) study (Ozkaynak et al 1996), personal exposure measurements were recorded for 48 consecutive days (September 22 to November 9, 1990) on some of 178 selected residents of Riverside, California. At most, 2 participants were monitored on any given day.
2. In the Harvard study of patients with chronic obstructive pulmonary disease (COPD) (Rojas-Bracho et al 1998), personal exposure measurements were made during the winter of 1996 and the winter and summer of 1997 for a total of 114 nonconsecutive days on 18 COPD patients living in Boston. At most, 4 participants were monitored on any given day.
3. In the Netherlands study of adults (NETH-A) (Janssen et al 1998), personal exposure measurements were made during the fall and winter of 1994, for a total of

Table 1. Study Characteristics, Averages of Personal and Outdoor PM_{10} Concentrations, and Results of Regression Model for Daily Personal PM_{10} Exposures Versus Ambient (or Outdoor) Concentrations

	Study				
Study Characteristics	PTEAM	COPD	NETH-A	NETH-C	THEES
Number of sampling days	49	114	43	26	14
Number of people	178	18	37	45	14
Number of daily observations	2–8	1–4	1–12	5–15	11–14
Mean personal exposure (μg/m^3)	110.58	37.29	60.38	104.55	85.85
Mean ambient exposure (μg/m^3)	84.51	23.19	40.82	38.20	60.05
Intercept[a] (α_0^s)	59.95 ± 4.1	29.92 ± 2.6	34.17 ± 2.19	85.08 ± 2.88	57.69 ± 8.31
Slope[a] (α_1^s)	0.60 ± 0.01	0.33 ± 0.12	0.72 ± 0.13	0.48 ± 0.13	0.45 ± 0.19

[a] Estimated regression coefficient ± SD.

43 nonconsecutive days on 37 nonsmoking adults living in Amsterdam, The Netherlands. At most, 12 participants were monitored on the same day.

4. In the Netherlands study of children (NETH-C) (Janssen et al 1997), personal exposure measurements were made in the spring of 1994 and the subsequent fall of 1995, for a total of 45 nonconsecutive days on 45 children from 4 schools in Amsterdam, The Netherlands. Four to eight repeated measurements of personal PM_{10} concentrations were obtained for each child.
5. Finally, in the Total Human Environmental Exposure Study (THEES) (Lioy et al 1990; Mage and Buckley 1995), PM_{10} personal samples were taken in the winter of 1988 for a total of 14 consecutive days on 14 nonsmoking individuals in Phillipsburg, New Jersey. A total of 11 to 14 repeated measurements of personal PM_{10} concentrations were taken for each subject.

For the first 4 studies, each had a single outdoor monitoring site recording ambient PM_{10} concentrations for the entire period, while there were 4 monitors recording data in THEES. Because the variation in the outdoor PM_{10} concentrations among outdoor sites on a single day was small in THEES, the daily data have been aggregated by taking their daily mean.

Table 1 summarizes the 5 studies showing the sampling schemes, the averages of personal and outdoor PM_{10} concentrations, and the results of a linear regression model for daily personal PM_{10} exposures versus ambient (or outdoor) concentrations. In all of the studies, mean personal exposures to PM_{10} were usually greater than ambient concentrations, probably because of the influence of indoor sources such as smoking, cooking, and dust on exposures to particles.

METHODS

LOG-LINEAR REGRESSION OF MORTALITY ON MEASURED AMBIENT POLLUTION CONCENTRATIONS

In this subsection, we describe time-series models for investigating the effects of measured ambient concentrations of PM_{10} on mortality risk. To reestablish notation, Y_t, X_t, and Z_t are the observed mortality, average personal exposure to PM_{10}, and the measured ambient PM_{10} concentration on day t, respectively. β_z is the log relative rate of mortality associated with a unit increase in Z_t, the ambient concentration of PM_{10}. Here, β_z is the target of estimation. In the general framework for measurement error, the same model is used with X_t replacing Z_t, thus giving inferences about β_x for comparison with those on β_z.

To estimate β_z, we use a generalized additive model (Hastie and Tibshirani 1990) with log link and Poisson error. We account for seasonal and longer-term fluctuations in mortality and temperature that can confound the pollution effects (Samet et al 1995; Kelsall et al 1997) using spline functions of calendar time and temperature, respectively. We consider 3 age groups: younger than 65 years old, 65 to 75 years, and older than 75 years. More specifically:

$$\begin{aligned}\log \mu_t = {} & \beta_0 I_{<65} + \beta_1 I_{65-75} + \beta_2 I_{>75} + \beta_z Z_t \\ & + \beta_{dow}\mathrm{DOW} + S_1(\mathrm{time}, 7/\mathrm{year}) \\ & + S_2(\mathrm{temp}_0, 6) + S_3(\mathrm{temp}_{1-3}, 6) \\ & + S_4(\mathrm{dew}_0, 3) + S_5(\mathrm{dew}_{1-3}, 3) \\ & + \text{smooth function of time (8 } df\text{) for age group} \end{aligned} \quad (1)$$

where β_0, β_1, and β_2 are the age-specific intercepts and *DOW* are indicator variables for day of week.

Smooth functions of calendar time S (time, λ) are included for each city to protect the estimate of pollution log relative rate $\hat{\beta}_Z$ from confounding by longer-term trends due to changes in overall health conditions, sizes and characteristics of populations, seasonality, and influenza epidemics, and to account for any additional temporal correlation in the count time series. That is, we estimate the pollution effect using only shorter-term variations in mortality and air pollution. Here, λ is *df* for the spline, which can be prespecified based on epidemiologic knowledge of the time scale of the possible confounder. For example, Dominici and colleagues (1999b) set $\lambda = 7$ *df* per year so that little information from time scales longer than approximately 2 months is considered in estimating β_z. To control for weather, we fit smooth functions of the same day temperature (temp_0), average temperature for the 3 previous days (temp_{1-3}), each with 6 *df*, and the analogous functions for dew point (dew_0, dew_{1-3}), each also with 3 *df*. We also control for age-specific longer-term temporal variations in mortality, adding a separate smooth function of time with 2 *df* per year, for each age-group contrast, to allow for different long-term trends across the 3 age groups. We impose a finite-dimensional parameter space on $S(x, \lambda)$ by restricting the choice of the smooth functions to the space spanned by a finite set of natural cubic splines on a fixed grid of knots. Each function $S(x, \lambda)$ can therefore be rewritten in the linear fashion $\sum_{j=1}^{q} B_j(x)\gamma_j$ for some values $\gamma_1 \ldots \gamma_q$ (Green and Silverman 1994).

Model (1) can be rewritten as follows:

$$\log \mu_t = W_{zt}\theta_z + B_t\gamma_z \quad (2)$$

where $W_{zt} = [Z_t, I_{<65}, I_{65–75}, I_{>75}, DOW]$ and $\theta_z = [\beta_z, \beta_0, \beta_1, \beta_2, \beta_{dow}]$. Here, B_t is the tth row of the design matrix for the cubic splines and γ_t is the corresponding vector of coefficients. Kelsall and colleagues (1997) and more recently Dominici and colleagues (2000) have used this model to estimate β_z in Philadelphia and in the 20 largest US locations, respectively.

GENERAL FRAMEWORK FOR MEASUREMENT ERROR

We begin our discussion of measurement error by reviewing the conceptual framework in the log-linear regression used for time-series studies of particulate air pollution and mortality (Zeger et al 1999). The discussion below is based on the premise that the ideal investigation of the health effects of air pollution would be conducted at the individual level with measurement of personal exposures to the pollutants. Exposure and mortality are only available after aggregation to a municipal level, however; pollution data from indoor monitoring are not available in the analysis.

To estimate the log relative rate of mortality for increases of personal exposure to PM, we would ideally consider a model

$$\mu_{it} = \mu_{0it}\exp(X_{it}\beta_x + \textit{confounders}) \tag{3}$$

where μ_{it} is the individual's risk of dying on day t; μ_{0it} is the baseline risk for person i when all predictors are 0; X_{it} is personal exposure that day; β_x is the log relative rate of scientific interest, and *confounders* represents all terms in the log-linear regression model (2) except $\beta_z Z_t$. By (3), the expected total numbers of deaths μ_t in a community of size n_t is

$$\mu_t = \sum_i \mu_{0it}\exp(X_{it}\beta_x + \textit{confounders}). \tag{4}$$

In analyzing population-level data on mortality and air pollution, we have previously used model (2), which estimates β_z, the log relative rate of mortality for increases of ambient concentrations. If the regression coefficient β_x in model (3) is the target of inference, how closely do estimates of β_z from model (2) approximate β_x?

Because we observe Z_t and not X_{it}, we can write

$$X_{it} = (X_{it} - X_t) + (X_t - Z_t) + Z_t,$$

where $X_{it} - X_t$ is the difference between an individual's exposure X_{it} and the population-average personal exposure X_t, and $X_t - Z_t$ is the difference between average personal exposure X_t and the measured ambient level Z_t. If we substitute this decomposition into equation (4), and limit attention to first-order terms in a Taylor series expansion, the expected number of deaths on day t can be approximated:

$$\log \mu_t \cong \log(n_t\mu_{0t}) + \beta_x(X_t^{(w)} - X_t) + \beta_x(X_t - Z_t) + \beta_x Z_t + \textit{confounders}. \tag{5}$$

Here $n_t\mu_{0t}$ denotes the total baseline risk in the population of size n_t, and $X_t^{(w)}$ is a risk-weighted personal exposure given by $X_t^{(w)} = \sum_i \mu_{0it}X_{it}/\sum_i \mu_{0it}$, and X_t denotes the unweighted average exposure. The first term, $n_t\mu_{0t}$, will vary slowly over time because the population and its health behaviors cannot change suddenly. The second term, $\beta_x(X_t^{(w)} - X_t)$, represents the effect of having aggregated rather than individual data in an ecological analysis such as ours. It is likely to be small and also to vary slowly over time t unless high-risk individuals change their exposures to pollutants over the shorter term in response to pollution levels. This, however, is unlikely.

The third error term, $\beta_x(X_t - Z_t)$, has the greatest potential to introduce bias in the estimate $\hat{\beta}_z$ and hence it is the focus of our analysis in the next subsection. See Zeger and colleagues (1999) for a more detailed analysis of sources of measurement error in estimates of air pollution exposure.

Because day-to-day variation in mortality attributable to air pollution is on the order of 1% or 2% (Kelsall et al 1997), the linear approximation above is reasonable. More detailed analysis of higher-order terms in the Taylor series expansion is unlikely to alter the inferences below but may offer some additional insights for studies of morbidity where attributable risks could be an order of magnitude higher.

MODELING MORTALITY AS A FUNCTION OF AVERAGE PERSONAL EXPOSURE

The regression methodology used to estimate the mortality-pollution association is extended here by accounting for measurement error in PM_{10}. We assume that the ambient measure of PM_{10}, denoted by Z_t in (2), is a surrogate measure of the average personal concentration X_t (Carroll et al 1995); that is, Y_t is conditionally independent of Z_t given X_t. Our strategy can be described in 2 steps. First, we model the mortality/personal exposure association in a particular location (as in the Baltimore example below) by the generalized additive model (2) with X_t, the unknown average personal exposure to PM_{10}, as the key predictor rather than Z_t. We denote the observed and missing predictors for Baltimore by Z_t^B and X_t^B. We then use 5 additional data sources on average personal

exposure and ambient concentrations to estimate X_t^B and β_x. Let $S = \{1,2,3,4,5,B\}$ be the label-set of data sources. The available data are the mortality counts in Baltimore Y_t^B, the ambient PM_{10} measures for all the locations Z_t^s, $s \in S$, and the average personal PM_{10} exposures for all the locations X_t^s except for Baltimore. We model these data in 2 parts.

The model for Y_t^B given X_t^B is specified by

$$Y_t^B \mid \mu_t \sim \text{Poisson}(\mu_t),\ t = 1, \ldots T$$
$$\log \mu_t = W_{xt}\theta_x + B_t\gamma_x \quad (6)$$

where $\mu_t = \mu_t(X_t^B)$ since $W_{xt} = [X_t^B, I_{<65t}, I_{65-75t}, I_{>75t}, DOW]$ and $\theta_x = [\beta_x, \beta_0, \beta_1, \beta_2, \beta_{dow}]$.

The hierarchical model for X_t^s given Z_t^s is then specified by

$$X_t^s = \alpha_0^s + \alpha_1^s Z_t^s + \epsilon_t, \epsilon_t \sim N(0,\sigma_x^2),\ s \in S$$
$$\alpha_0^s \mid \alpha_0 \sim N(\alpha_0, \tau_0^2),\ s \in S \quad (7)$$
$$\alpha_1^s \mid \alpha_1 \sim N(\alpha_1, \tau_1^2),\ s \in S.$$

The slope α_1^s measures the change in personal exposure per unit change in measured ambient concentration at location s. If there is relatively little measured error in the ambient levels, the slope can be also interpreted as the fraction of exposure from ambient sources that occurs either outdoors or through penetration of ambient pollution indoors. The variances σ_x^2, τ_0^2, and τ_1^2 represent the error in the estimated regression of X_t^s on Z_t^s, and the variability of the regression coefficients α_0^s and α_1^s across studies, respectively.

Under this simple linear model, the intercept α_0^s represents personal exposure to particles that does not derive from external sources but arises from particle clouds generated by personal activities or unmeasured microenvironments. Because few of the available data sets report personal and ambient data on several people, we assume that α_0^s and α_1^s are constant across subjects within a region. As more data on personal exposures and ambient concentrations become available, however, this model can be readily extended to a longitudinal regression model with subject-specific slopes. Doing so in this application has little effect on the results because the number of subjects per region is much larger than the number of regions.

PRIOR DISTRIBUTIONS AND COMPUTATION

For a Bayesian analysis with this model, we must specify prior distributions for all unknown parameters. An attractive and practical approach in a hierarchical model is to specify dispersed but proper prior distributions and then supplement the baseline analysis with additional sensitivity analyses. A priori, the unknown parameters were assumed to be independent so that the joint prior is the product of the marginal of each parameter. For the overall regression parameters α_0, α_1, and for the vector of the log relative rates θ_x, we use normal distributions, and for the variance parameters τ_0^2, τ_1^2, and σ_x^2 we use inverse gamma distributions. Prior means and 95% intervals of all the unknown parameters are summarized in Table 2.

Since the data on the association between X_t and Z_t are available from only 5 studies, the prior distribution for τ_0^2 and τ_1^2—that is, the variance across cities in the study-specific intercepts and slopes—has a substantial impact on our posterior inference. The prior specification for these parameters has been selected to allow the personal exposure to particles that does not derive from external sources, α_0^s, and the fraction of personal exposure that derives from measured ambient concentrations, α_1^s, to range across a large set of reasonable values. Later we supplement this baseline analysis with additional sensitivity analysis. Under this model, samples from the posterior distribution of the unknown parameters can be drawn by implementing a block Gibbs sampler (Gelfand and Smith 1990) with Metropolis steps to draw from the full conditional distributions of θ_z, γ_z under model (2) and θ_x, γ_x, and X_t^B under model (3). For θ_x and γ_x, we use a random walk proposal where we generate each component vector from a normal distribution centered at the current value of the parameter and with variance obtained from the output of a generalized additive model with log link Poisson error.

Table 2. Means and 95% Intervals A Priori and A Posteriori of Unknown Parameters

Parameters	A Priori	A Posteriori
α_0	50 (–16, 114)	51.6 (32.05, 71.45)
α_1	0.5 (–1.5, 2.55)	0.53 (0.21, 0.86)
α_0^B	$\alpha_0(\alpha_0 - 1.96\tau_0, \alpha_0 + 1.96\tau_0)$	49 (22.71, 73.32)
α_1^B	$\alpha_1(\alpha_1 - 1.96\tau_1, \alpha_1 + 1.96\tau_1)$	0.60 (0.13, 1.13)
σ_x	20 (12, 42)	18.71 (17.17, 20.62)
τ_0	20 (12, 42)	20.64 (13.53, 32.39)
τ_1	0.43 (0.27, 0.85)	0.35 (0.23, 0.55)
β_x	0 (–10, 10)	1.4 (0.24, 2.88)
β_z	0 (–10, 10)	0.9 (0.67, 1.12)

The same strategy is used to sample θ_z and γ_z. For X_t^B we use an independent proposal equal to the normal prior distribution specified by equation (7). Because we will gain little information about X_t^B from Y_t^B, this proposal choice leads to an efficient strategy.

When γ_z and γ_x are high dimensional, the computations needed for implementing a full Bayesian approach—that is, to draw from the joint posterior distributions of θ_z and γ_z and then to integrate over the γ_z to obtain the marginal posterior distributions of the θ_z—are feasible but extremely laborious. The computation becomes even more intensive if we want to make inferences on β_x, taking into account the uncertainty due to the lack of knowledge of X_t^B.

To ease the computational burden, we have approximated the posterior distributions of θ_z and θ_x by assuming $\gamma_z = \hat{\gamma}_z$ under model (2) and $\gamma_x = \hat{\gamma}_x$ under model (3), where $\hat{\gamma}_z$ and $\hat{\gamma}_x$ are the maximum likelihood estimates obtained under the generalized additive model with log link and Poisson error. With these assumptions, the difference between the marginal posterior distributions of θ_z and θ_x obtained under the full Bayesian approach and its approximation (where $\gamma_z = \hat{\gamma}$) is small and leads to no meaningful differences in the inferences on all other unknown parameters.

ANALYSIS OF BALTIMORE DATA

In this subsection we analyze air pollution data from Baltimore, Maryland, using additional information obtained from the 5 epidemiologic studies on personal and outdoor exposures to PM_{10}. We start by presenting a non-Bayesian nonhierarchical method based on a regression calibration approach. We then use the Bayesian hierarchical model introduced previously and compare the results obtained under the two approaches.

REGRESSION CALIBRATION APPROACH

An alternative non-Bayesian method uses the following 2-stage regression calibration. At the first stage, we fit linear regression models to the 5 validation data sets. We then estimate the overall intercept and slope $\hat{\alpha}_0$, $\hat{\alpha}_1$ by using a weighted average for a random effect model—that is, we average the study-specific intercepts and slopes in Table 1 with modified weights to take into account the variability of coefficients among the studies (DerSimonian and Laird 1986).

We obtain $\hat{\alpha}_0 = 53.18$ and $\hat{\alpha}_1 = 0.53$. At the second stage, we estimate the average personal exposure time series for the location of interests as $\hat{X}_t^B = \hat{\alpha}_0 + \hat{\alpha}_1 Z_t^B$, and we use $\hat{X}_t^B$ instead of Z_t^B in the generalized additive model (6). For Baltimore, we obtain $\hat{\beta}_x = 1.67$ with 95% confidence interval (CI) equal to (0.26, 3.21).

HIERARCHICAL BAYESIAN ANALYSIS

In this subsection, we analyze the Baltimore data by using the hierarchical Bayesian model (6), (7), presented previously. Model fitting is performed using MCMC methods. Because of the lack of measurements on personal exposure in Baltimore, the sampled values of our parameter of interest β_x are likely to be autocorrelated, resulting in a chain that slowly converges to the posterior distribution. To assess convergence, we used the Gelman and Rubin (1992) diagnostic in Convergence Diagnostics and Output Analysis (CODA) (Best et al 1995). We run 5 parallel chains starting from 5 overdispersed sets of initial values. We run the chains for 100,000 iterations and save every tenth value. The value of the shrinkage factor for β_x has a median of 1.01 and 97.5% upper CI equal to 1.03, suggesting that the 5 chains mix well and may be assumed to arise from the desired marginal posterior distribution. The acceptance probabilities for the Metropolis steps for θ_x (and for θ_z in model [2]) and for X_t^B were roughly equal to 0.4 and 0.9, respectively. The high acceptance rate for X_t^B indicates that the mortality data Y_t^B is only slightly informative in the estimate of X_t^B.

Figure 2 gives a summary of posterior inferences on study-specific regressions for personal concentrations

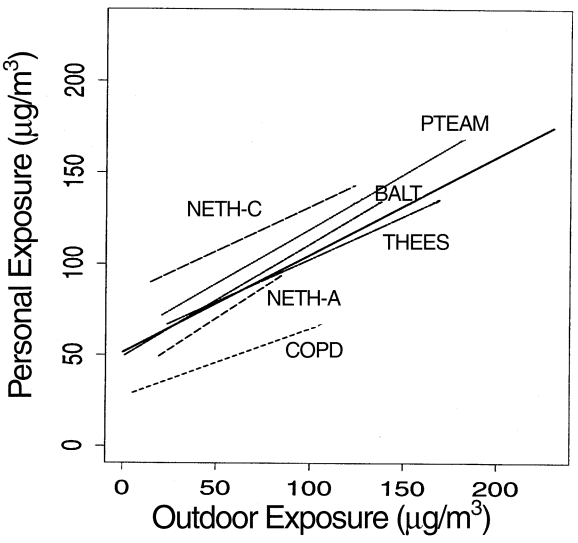

Figure 2. Summary of posterior inferences on study-specific regressions for personal concentration versus ambient measures of PM_{10}. The picture summarizes the point inference of regression coefficients. Each line is the regression line given by the posterior mean of the overall parameters. Each regression line is plotted at the range of the study-specific outdoor exposures. The regression line for the Netherlands studies (children) is above the diagonal, showing an average personal exposure always higher than the outdoor exposure.

versus ambient measures of PM_{10}. Each line is the regression defined by the posterior means of the study-specific parameters, α_0^s, and α_1^s. The thicker line is the regression line defined by the posterior means of the overall parameters α_0 and α_1. Each regression line is plotted on the range of the study-specific outdoor exposures. The regression line for the Netherlands studies (children) is above the diagonal, showing an average personal exposure that is always higher than the outdoor concentrations.

Figure 3 shows the posterior distributions of the study-specific intercepts and slopes, respectively. At the far right, the posterior distributions of the corresponding overall parameter values α_0 and α_1 are pictured. The boxplots for Baltimore are wider because they incorporate the uncertainty from lack of knowledge about personal concentrations for that location.

Overall, we find that a unit change in ambient concentration is associated with a 0.53 increase in average personal exposure, and that the intercept has a posterior mean of 51.6 $\mu g/m^3$. These results are very close to the estimated $\hat{\alpha}_0$ and $\hat{\alpha}_1$ of the regression calibration model. The posterior mean and 95% interval of the regression standard error σ_x in the estimated regression is 18.71. Posterior means of α are all close to the estimated regression coefficients summarized in Table 1. In addition to the posterior distributions of the regression coefficients for the 5 validation studies, the Bayesian hierarchical model provides estimates of the regression coefficients α_0^B and α_1^B for Baltimore. The estimated marginal posterior means are 49 and 0.65, respectively. Because the hierarchical model allows for heterogeneity across studies of α, these estimates differ slightly from the overall regression coefficients $\hat{\alpha}_0$, $\hat{\alpha}_1$ estimated under the regression calibration. The variability of the regression coefficients (intercept and slopes) between data sources is captured by τ_0 and τ_1. The estimated marginal posterior means are 20.64 and 0.35, respectively. Means and 95% regions a priori and a posteriori of all the unknown parameters are summarized in Table 2.

Figure 4 displays the estimate of the nonlinear adjustment for the current day temperature. The vertical scale can be interpreted as the relative risk of mortality as a function of temperature. The dotted lines are the 95% confidence interval bands at each given temperature. As expected, the highest daily mortality mean occurs at the most extreme temperatures; the minimum level of daily mortality is reached at 69°F.

Finally, Figure 5 shows the posterior distributions of log relative rates (percentage increase in mortality for 10 $\mu g/m^3$ increase in PM_{10} exposure) of mortality from ambient exposures β_z and from total personal exposure β_x. The

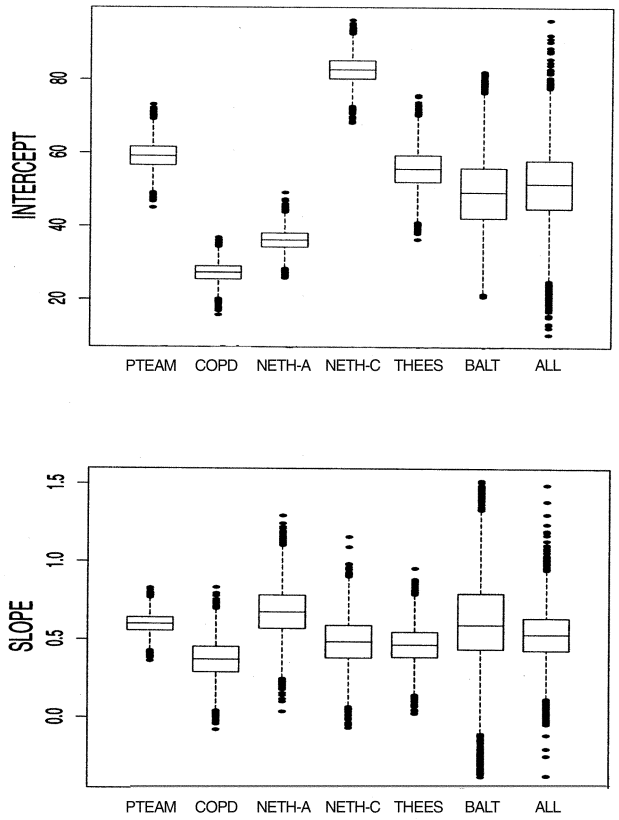

Figure 3. Boxplots of samples from the posterior distributions of the study-specific intercepts and slopes. At the far right are the posterior distributions of the corresponding overall effects. The boxplots for Baltimore are wider because they incorporate uncertainty due to the absence of personal concentrations for that location.

posterior means and interquartile range (IQR) of β_z and β_x are 0.9, (0.67, 1.12) and 1.4, (0.24, 2.88), respectively.

Note that measurement error tends to bias the results toward zero and that the IQR of β_x is larger than the IQR of β_z. The posterior standard deviation of β_x, which is equal to 0.74, is larger than the posterior standard deviation of β_z, 0.34, for 2 reasons: (1) even if we know with certainty the correction factor α_1^B, the intercept α_0^B, and the average personal exposure time series in Baltimore X_t^B, then the posterior standard deviation of β_x would be roughly equal to the posterior standard deviation of β_z times $1/\alpha_1^B$; (2) the IQR of β_x also incorporates the uncertainty from estimating the correction factor and not having the average personal exposure time series in Baltimore, Maryland.

The estimate of β_x is slightly larger under the regression calibration model than under the Bayesian hierarchical model, with a substantial overlap of the 2 IQRs. The hierarchical model on the study-specific regression coefficients, α_0^s and α_1^s, and the estimation approach for the missing personal exposure to PM_{10} for Baltimore, X_t^B, contribute to

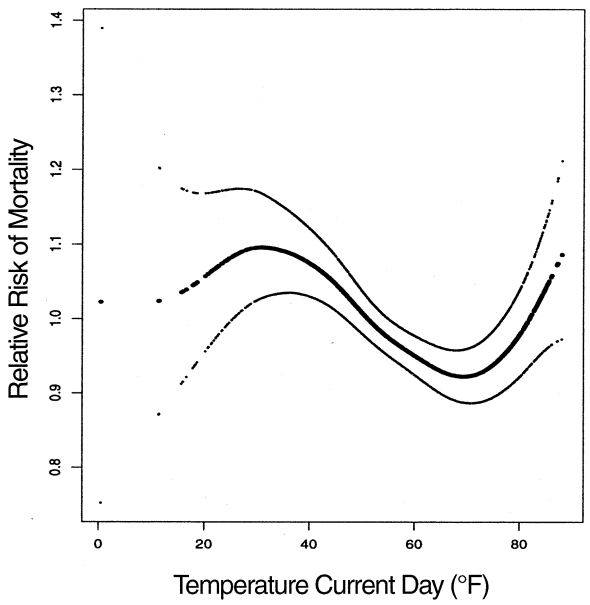

Figure 4. Nonlinear adjustment for the current day temperature. The vertical scale can be interpreted as the adjusted daily mortality mean as a function of the temperature. The thinnest lines are the 95% posterior bands at each given temperature. As expected, the highest daily mortality mean occurs at the most extreme temperatures; the minimum level of daily mortality is reached at 69°F.

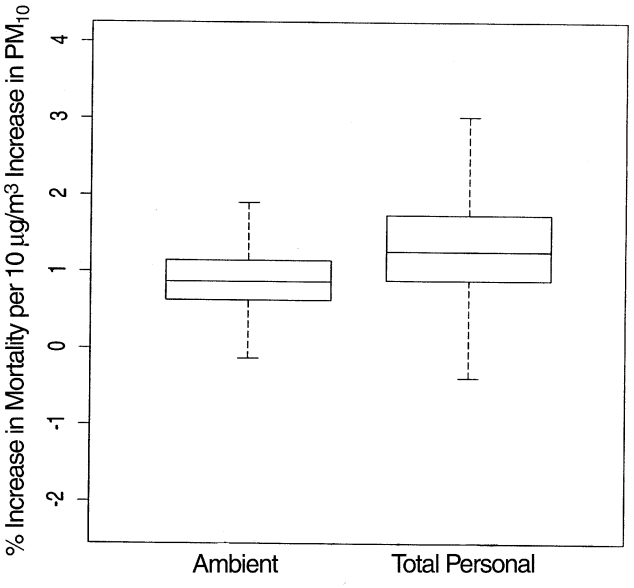

Figure 5: Comparison between the relative rates (% increase in mortality for 10 µg/m^3 increase of PM_{10} exposure), obtained by fitting the semiparametric Poisson model having as predictors the ambient concentrations and the average personal exposures, respectively. The boxplots represent the posterior distributions of the 2 relative rates (βz, βx). The posterior means and IQR of βz and βx are 0.47, (0.1, 0.8) and 0.80, (0.33, 1.60), respectively.

this difference. First, in the regression calibration, we assume that α_0^B and α_1^B equal the overall coefficients $\hat{\alpha}_0$ and $\hat{\alpha}_1$. The hierarchical model incorporates variability in the α across the studies' regressions and thus allows α_0^B and α_1^B to be different from the population parameters. Second, in the regression calibration, we estimate X_t^B with a plug-in procedure without accounting for the information from Y_t^B. Under a Bayesian framework, X_t^B is considered to be an unknown parameter and therefore we estimate its marginal posterior distribution, $p(X_t^B \mid Y_t^B, Z_t^B, \text{data})$, using all available data.

In summary, the Bayesian model is a more conservative approach because it takes into account key sources of uncertainty in the log relative rate $\hat{\beta}_x$, which results from estimating the regression coefficients $\hat{\alpha}_0$,$\hat{\alpha}_1$, as well as the average personal exposure X_t^B. A strength of the Bayesian model is that it easily provides a quantitative assessment of the variability across studies in the personal–ambient relationships by estimating the posterior distribution of the variance parameters τ_0^2 and τ_1^2. In addition, the 2-stage regression calibration approach does not easily lend itself to generating ranking probabilities as, for example, $P(\beta_x \geq \beta_z \mid \text{data})$. These advantages come at the cost of increased computational complexity and of the introduction of prior information that necessitates further sensitivity analyses.

MODEL CHECKING AND SENSITIVITY ANALYSIS

Our strategy for investigating the impact of the model assumptions and the prior distribution on our results is based on inspecting posterior summaries of β_x under the following scenarios for departure from the hierarchical model (7), and from the baseline prior distribution.

In the measurement error model (7), one of the most influential assumptions is likely to be additive Gaussian error. Hence, we consider the following alternative scenario:

1. in (7), we assume a log normal distribution with multiplicative error, that is,

$$\log(X_t^S) = \lambda_0^S + \lambda_1^S \log(Z_t^S) + N(0,\delta^2).$$

The 2 most influential prior parameters in the sensitivity analysis are the variances across cities in the intercept and slope of the regression of X_t on Z_t, since data are available for only 5 studies. The alternative scenarios for the sensitivity analyses to the prior distribution are:

2. in the baseline, with the assumption that the variability across studies in the personal exposure to ambient is believed to be 2 times higher: $B_1 = 0.5 \times 2$;

3. in the baseline, with the assumption that the variability across studies in the personal exposure to ambient is believed to be 2 times smaller: $B_1 = 0.5/2$;
4. in the baseline, with the assumption that the variability across studies in the indoor exposures is believed to be 2 times higher: $B_0 = 1{,}000 \times 2$;
5. in the baseline, with the assumption that the variability across studies in the indoor exposures is believed to be 2 times smaller: $B_0 = 1{,}000/2$; and
6. in the baseline, with the assumption that the prior mean of β_x is believed to be equal to the regression calibration estimate: $E[\beta_x] = 1.67$.

Figure 6 shows the marginal posterior distributions of β_x and the posterior probabilities that $\beta_x > \beta_z$ under our baseline model and under the 6 alternative scenarios. The posterior probabilities of $\beta_x \geq \beta_z$, posterior means, and IQR of β_x are quite stable under all scenarios, revealing that the hierarchical model is not very sensitive to model assumptions and the prior distributions.

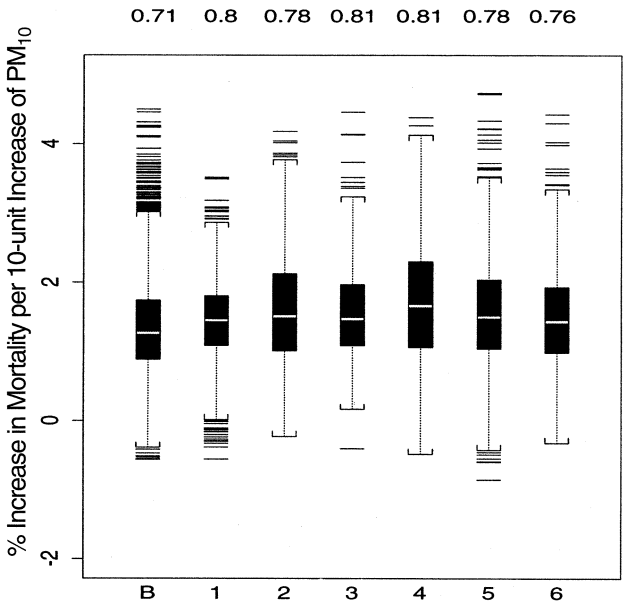

Figure 6. Posterior distributions of β_x under the 6 alternative scenarios. At the top are shown the posterior probabilities that $\beta_x \geq \beta_z$. B = baseline.

DISCUSSION

Findings of air pollution–mortality studies figure prominently in the development of environmental public policy. Hence, a better understanding of the consequences of the measurement error in air pollution studies is needed (Thomas et al 1993; Zidek et al 1998; Zeger et al 1999). We have developed a hierarchical measurement error model to combine information about the relationship of mortality to ambient air quality measures and the association of ambient concentrations and personal exposures to particulate pollution exposure. Our model combines 2 sources of information to estimate the coefficient β_x, which measures the relative increase in mortality per unit increase in personal exposure to PM_{10}. Like every statistical model, this one is an approximation of reality, designed to make effective use of the available information about β_x.

We have estimated both β_z and β_x in Baltimore. The parameter β_x describes the effect of personal exposure to particulates on the risk of mortality, as discussed in section 3. It is of direct biologic and etiologic interest. If ambient concentrations Z_t are measured precisely, β_z represents the expected proportional change in mortality in a population given a unit reduction in ambient air pollution levels. Since, in most countries, governments attempt to regulate ambient concentrations rather than personal exposures, β_z is of direct interest to regulators. In the models presented here, β_z can be estimated directly from the available data Y_t and Z_t for a city such as Baltimore. We depend on auxiliary information and modeling assumptions to estimate β_x. Nevertheless, it is useful to have an estimate of the bias arising from using ambient levels rather than personal exposures, even if that estimate is dependent on assumptions. Heterogeneity in this bias across locations is likely, and it is relevant to assessing the external validity of regression findings in a particular location.

It is also useful to have an estimate of the effect of personal exposure to ambient measurements on the risk of mortality. Although such an estimate must take into account the different sources of emissions, we can give a partial answer for Baltimore by inspecting the posterior distribution of $\beta_x \times \alpha_1^B$. The posterior mean and IQR of this parameter is 0.79, (0.56, 1.02), which is close to our direct estimate of the posterior distribution of β_z.

The strengths of our analysis are that

1. it extends regression calibration, giving a more conservative result because it takes into account mortality data Y_t^B and all the key sources of uncertainty in the log relative rate $\hat{\beta}_x$;
2. it easily provides a quantitative assessment of the variability across studies in the personal–ambient relationships;
3. it lends itself to generating ranking probabilities as, for example, $P(\beta_x \geq \beta_z \mid \text{data})$; and
4. the results do not appear sensitive to alternative measurement error model assumptions and specifications of the prior distribution within a reasonable range.

One important limitation of the current formulation is the modeling of the association between ambient and average personal exposure. Given extensive daily ambient exposure data, a time-series model of X_{it} on Z_t would be more appropriate than the simple linear models we used. In the PTEAM data, we do have 48 consecutive days of both ambient and average personal exposure levels and have examined whether their association is identical at all time scales, as assumed by our linear model. In fact, we see some evidence that the association is stronger at longer time scales, suggesting that a time-series model might improve on the analysis presented here. But the PTEAM study is the only one with time-series data and even these data are limited in duration, making additional exploratory analysis difficult. Extensions of our model to allow for time-series structure in the exposure component can be easily implemented, given time-series data at several sites. A desirable extension of our model would be to allow heterogeneity across studies of the regression variances. In Baltimore, however, because X_t^B is unknown, such variance is not identifiable.

Another limitation is that our measurement error analysis is for a single pollutant, PM_{10}. An important and needed extension would be the case of multiple pollutants—for example, PM_{10}, NO_2, SO_2, and CO, all combustion products—which are correlated and simultaneously measured with correlated errors. Zidek and colleagues (1996) and Zeger and colleagues (1999) have investigated the effects of measurement error in regression analyses when explanatory variables are correlated and measured with errors. Currently, we are unaware of empirical studies that have measured multiple pollutants for individuals and at central sites simultaneously. Our model extends naturally to the multipollutant case when such data become available in the near future.

ACKNOWLEDGMENTS

The authors are grateful to Drs H Ozkaynak, J Xue, J Spengler, L Wallace, E Pellizzari, P Jenkins, L Rojas-Bracho, H Suh, P Koutrakis, NAH Janssen, G Hoek, B Brunekreef, H Harssema, I Mensink, A Zuidof, PJ Lioy, JM Waldman, T Buckley, J Butler, and T Pietarinen, who kindly provided their data for the 5 studies on outdoor and personal exposures to particulate matter. The authors thank Drs Giovanni Parmigiani and Duncan Thomas for encouragement and suggestions and are also grateful to the reviewers and associate editor for their comments.

REFERENCES

Bascom R, Bromberg PA, Costa DA, Devlin R, Dockery DW, Frampton MW, Lambert W, Samet JM, Speizer FE, Utell M. 1996a. Health effects of outdoor air pollution. Part 1. Am J Respir Crit Care Med 153:3–50.

Bascom R, Bromberg PA, Costa DA, Devlin R, Dockery DW, Frampton MW, Lambert W, Samet JM, Speizer FE, Utell M. 1996b. Health effects of outdoor air pollution. Part 2. Am J Respir Crit Care Med 153:477–498.

Armstrong B. 1985. Measurement error in the generalized linear model. Commun Stat, Part B: Simul Comput 14: 529–544.

Best NG, Cowles MK, Vines K. 1995. CODA: Convergence Diagnostics and Output Analysis Software for Gibbs Sampling Output, Version 0.30. Technical Report. MRC Biostatistics Unit, University of Cambridge, Cambridge, England.

Carroll RJ, Ruppert D, Stefanski L. 1995. Measurement Error in Nonlinear Models. Chapman & Hall, New York NY.

Carroll RJ, Stefanski LA. 1990. Approximate quasi-likelihood estimation in models with surrogate predictors. J Am Stat Assoc 85:652–663.

Clayton D. 1991. Models for the longitudinal analysis of cohort and case-control studies with inaccurately measured exposures. In: Statistical Models for Longitudinal Studies of Health. Oxford University Press, New York NY.

Dellaportas P, Stephens DA. 1997. Bayesian analysis of errors-in-variables regression models. Biometrics 51:1085–1095.

DerSimonian R, Laird N. 1986. Meta-analysis in clinical trials. Control Clin Trials 7:177–188.

Dockery D, Pope C. 1994. Acute respiratory effects of particulate air pollution. Annu Rev Public Health 15:107–132.

Dominici F. 2000. Combining contingency tables with missing dimensions. Biometrics. In press.

Dominici F, Parmigiani G, Reckhow KH, Wolpert RL. 1997. Combining information from related regressions. J Agric Biol Environ Stat 2:313–332.

Dominici F, Parmigiani G, Wolpert RL, Hasselblad V. 1999. Meta-analysis of migraine headache treatments: Combining information from heterogeneous designs. J Am Stat Assoc 94:16–28.

Dominici F, Samet J, Zeger S. 2000. Combining evidence on air pollution and daily mortality from the twenty

largest US cities: A hierarchical modeling strategy. J R Stat Soc, Series C. In press.

Fuller WA. 1987. Measurement Error Models. John Wiley & Sons, New York NY.

Gelfand AE, Smith AFM. 1990. Sampling-based approaches to calculating marginal densities. J Am Stat Assoc 85(410):398–409.

Gelman A, Rubin DB. 1992. Inference from iterative simulation using multiple sequences. Stat Sci 7:457–472.

Green PJ, Silverman BW. 1994. Nonparametric Regression and Generalized Linear Models: A Roughness Penalty Approach. Chapman & Hall, New York NY.

Hastie TJ, Tibshirani RJ, eds. 1990. Generalized Additive Models. Chapman & Hall, New York NY.

Janssen N, Hoek G, Brunekreef B, Harssema H, Mensink I, Zuidhof A. 1998. Personal sampling of particles in adults: Relation among personal, indoor, and outdoor air concentrations. Am J Epidemiol 147:537–544.

Janssen N, Hoek G, Harssema H, Brunekreef B. 1997. Childhood exposure to PM_{10}: Relation between personal, classroom, and outdoor concentrations. Occup Environ Med 54:1–7.

Kelsall J, Samet J, Zeger S. 1997. Air pollution and mortality in Philadelphia, 1974–1988. Am J Epidemiol 146:750–762.

Lindley DV, Smith AFM. 1972. Bayes estimates for the linear model (with discussion). J R Stat Soc 34:1–41.

Lioy P, Waldman J, Buckley T, Butler J, Pietarinen T. 1990. The personal, indoor, and outdoor concentrations of PM_{10} measured in an industrial community during the winter. Atmos Environ 24:57–66.

Mage T, Buckley TJ. 1995. The relationship between personal exposures and ambient concentrations of particulate matter. In: Air and Waste Management, 88th Annual Meeting, San Antonio TX.

Mallick BK, Gelfand AE. 1994. Errors-in-variables models: A Bayesian approach. Biometrika 81:237–245.

Morris CN, Normand S-L. 1992. Hierarchical models for combining information and for meta-analysis. In: Bayesian Statistics (Bernardo JM, Berger JO, Dawid AP, Smith AFM, eds), pp 321–344. Oxford University Press, Oxford, England.

Ozkaynak H, Xue J, Spengler J, Wallace L, Pellizzari E, Jenkins P. 1996. Personal exposure to airborne particles and metals: Results from the particle team study in Riverside, California. J Expo Anal Environ Epidemiol 6:57–78.

Prentice RL. 1982. Covariate measurement errors and parameter estimation in a failure time regression model. Biometrika 69:331–342.

Richardson S, Gilks WR. 1993. A Bayesian approach to measurement error problems in epidemiology using conditional independence models. Am J Epidemiol 138:430–442.

Robinson J, Nelson W. 1995. National Human Activity Pattern Survey Data Base. US Enviromental Protection Agency, Office of Air Quality Planning and Standards, Research Triangle Park NC.

Rojas-Bracho L, Suh H, Koutrakis P. 1998. $PM_{2.5}$ and PM_{10} Personal Exposures and Their Relationship to Outdoor Concentrations in a Group of COPD Participants Living in Boston. Technical Report. Harvard School of Public Health, Boston MA.

Rosner B, Willett WC, Spiegelman D. 1989. Correction of logistic regression relative risk estimates and confidence intervals for systematic within-person measurement error. Stat Med 8:1051–1069.

Samet JM, Zeger SL, Berhane K. 1995. The association of mortality and particulate air pollution. In: Particulate Air Pollution and Daily Mortality: Replication and Validation of Selected Studies, The Phase I Report of the Particle Epidemiology Evaluation Project. Health Effects Institute, Cambridge MA.

Schwartz J. 1995. Air pollution and daily mortality in Birmingham, Alabama. Am J Epidemiol 137:1136–1147.

Spiegelhalter D, Thomas A, Best N, Gilks W. 1994. BUGS: Bayesian Inference Using Gibbs Sampler, Version 0.30. MRC Biostatistics Unit, University of Cambridge, Cambridge, England.

Stefanski LA. 1985. The effects of measurement error on parameter estimation. Biometrika 72:583–592.

Tanner MA. 1991. Tools for Statistical Inference: Observed Data and Data Augmentation Methods, Vol 67, Lecture Notes in Statistics. Springer-Verlag, New York NY.

Tanner MA, Wong WH. 1987. The calculation of posterior distributions by data augmentation. J Am Stat Assoc 82:528–550.

Thomas D, Stram D, Dwyer J. 1993. Exposure measurement error: Influence on exposure-disease relationships and methods of corrections. Annu Rev Public Health 14:69–93.

Wallace L. 1996. Indoor particles: A review. J Air Waste Manage Assoc 46:98–126.

Zeger S, Thomas D, Dominici F, Cohen A, Schwartz J, Dockery D, Samet J. 1999. Exposure Measurement Error in Time-Series Studies of Air Pollution: Concepts and Consequences. Technical Report. Johns Hopkins University, Baltimore MD.

Zidek J, Wong H, Le N, Burnett R. 1996. Causality, measurement error and multicollinearity in epidemiology. Environmetrics 7:441–451.

Zidek J, Wong, L, Burnett R. 1998. Including structural measurement errors in the nonlinear regression analysis of clustered data. Can J Stat 26:537–548.

Section 3: Mortality Displacement–Resistant Estimates of Air Pollution Effects on Mortality

Scott L Zeger, Francesca Dominici, and Jonathan M Samet

ABSTRACT

A number of studies have recently shown an association between particle concentrations in outdoor air and daily mortality counts in urban locations. In the public health interpretation of this evidence, a key issue is whether the increased mortality associated with higher pollution levels is restricted to very frail persons for whom life expectancy is short in the absence of pollution. This possibility has been termed the *harvesting* or *mortality displacement* hypothesis. We present an approach to estimating the association between pollution and mortality from time-series data that is resistant to short-term mortality displacement. The method is based in the concept that mortality displacement alone creates associations only at shorter time scales. We use frequency domain log-linear regression (FDLLR)* to decompose the information about the pollution-mortality association into distinct time scales, and we then create mortality displacement–resistant estimates by excluding the short-term information that is affected by mortality displacement. We illustrate the methods with total suspended particles (TSP) and mortality counts from Philadelphia for 1974 to 1988. We show that the TSP-mortality association in Philadelphia is inconsistent with the mortality displacement–only hypothesis and that the mortality displacement–resistant estimates of the relative risk of mortality associated with TSP are actually larger, not smaller, than the ordinary estimates.

* A list of abbreviations and other terms appears on page 13.

The National Morbidity, Mortality and Air Pollution Study: Methods and Methodologic Issues, Part I of Health Effects Institute Research Report 94, includes an Investigators' Report, a Preface, a Commentary by the Health Review Committee, and an HEI Statement about the research project. Correspondence concerning this section may be addressed to Dr Scott L Zeger, Department of Biostatistics, Johns Hopkins School of Public Health, 615 North Wolfe Street, Baltimore MD 21205-2179.

Although this document was produced with partial funding by the United States Environmental Protection Agency under Assistance Award R824835 to the Health Effects Institute, it has not been subjected to the Agency's peer and administrative review and therefore may not necessarily reflect the views of the Agency, and no official endorsement by it should be inferred. The contents of this document also have not been reviewed by private party institutions, including those that support the Health Effects Institute; therefore, it may not reflect the views or policies of these parties, and no endorsement by them should be inferred.

INTRODUCTION

The acute effects on morbidity and mortality of extreme episodes of particulate air pollution have been well documented by the 1952 London fog and other air pollution disasters (Beaver 1953). More recently, acute health effects have been associated with fluctuations in particulate air pollution well within the US Environmental Protection Agency (EPA) standards (1995, 1996). Substantial evidence has accumulated over the last decade in support of associations of daily levels of air pollution with mortality counts and with measures of morbidity. Bascom and colleagues (1996a,b) and Dockery and Pope (1994) provide overviews of this new literature. Samet and colleagues (1995) have conducted reanalyses and critically evaluated the pioneering work by Schwartz and Dockery (1992) and largely confirm an acute association between mortality and particulate air pollution.

Nevertheless, uncertainty remains regarding the public health implications of these findings. First, controversy remains about whether a single constituent of air pollution is responsible for the increased mortality and morbidity or whether the adverse health effects are caused by combined actions of multiple pollutants (Samet et al 1997). Even if a single constituent of the mix of pollutants in urban air, such as small particles, is largely responsible for increasing morbidity and mortality, a second question arises: is the increase in mortality only among extremely frail individuals whose remaining life expectancy, in the absence of pollution, would be short? That is, are only a small number of total days of life lost from pollution, or are individuals dying who would otherwise have survived for substantial periods? The possibility that only extremely frail individuals die from exposure to air pollution has been termed the *harvesting* hypothesis (Schimmel and Murawski 1976), a phenomenon also referred to as *mortality displacement*.

One approach to investigating the possibility that only frail individuals are affected by air pollution uses a compartmental model. In the simplest frailty model, the death process is assumed to have 2 steps. First, an individual moves from a relatively healthier population into a very frail subgroup from which all mortality occurs. In the

second transition, persons in this frail pool die; the risk of dying increases at higher pollution levels. If the size of the pool of the frail persons is small, then the mean residency time in the frail condition is short, regardless of the exposure to pollution. In this case, mortality displacement can occur because persons in the frail pool have a short life expectancy in the absence of pollution. On the other hand, if the size of the frail population is large, the mean residency time in the frail state is relatively long, so that pollution-caused deaths substantially shorten life. More realistic extensions of this simple 2-compartment model—for example, to include 3 states: healthy, diseased, and highly frail—can be developed, but they would capture the mortality displacement principle in a similar way.

The evidence available in mortality time-series data to assess whether mortality displacement occurs is found in the pattern of mortality following days with a large number of deaths. If the frail subpopulation is small and if deaths can occur only from this pool, then the number of deaths on a day after a pollution event will be fewer than expected because the previous high mortality depletes the pool of at-risk frail individuals. Hence, we would expect a negative autocorrelation between the number of deaths on a day after a pollution episode and a day before that episode (Lipfert and Wyzga 1995).

A few investigators have developed statistical models using the ideas above to estimate the size of the frail population and the expected days of life lost due to exposure to air pollution. Smith and colleagues (1997) used a 2-compartment model as described above with the additional assumption that both the risk of becoming frail and the risk of death may depend on air pollution.

In this section, we take a different approach to the mortality displacement issue. Rather than attempting to estimate directly the degree of mortality displacement, we propose a class of estimators of the pollution-mortality association that is resistant to shorter-term mortality displacement. That is, we propose an approach that ignores the information in the time-series data in which short-term mortality displacement would influence the mortality-pollution association. With this approach, pollution relative risks are close to unity if the association is due to mortality displacement alone. The method is based on partitioning both the pollution and mortality time-series data into components with variation occurring at different time scales and then relying on the longer-term components to estimate the effect of pollution on mortality.

First, we propose a simple, 2-compartment model for mortality displacement and demonstrate that mortality displacement produces correlations between mortality and pollution data that are nonnegligible only at short time scales. Then we review briefly an approach to time-series modeling of the mortality–air pollution association that gives separate estimates of the pollution effect at different time scales. Kelsall and colleagues (1999) describe this FDLLR in detail. We next propose a mortality displacement–resistant estimator that sets aside the short-term associations that are subject to the influence of mortality displacement. Further, we apply this method to the analysis of particulate air pollution and mortality data from Philadelphia for the period 1974 to 1988, which were previously analyzed by several investigators (Kelsall et al 1997; Samet et al 1997).

SIMPLE MODEL FOR MORTALITY DISPLACEMENT

The simplest model that captures the mortality displacement phenomenon is based on the assumption that individuals in the general population transition into a very frail subgroup, and that death occurs only among this frail subpopulation. This idea can be implemented in the following difference equation $N_t = N_{t-1} - D_{t-1} + I_t$, where N_t is the size of the frail population, I_t is the number of new persons from the general population who become frail at the start of day t, and D_t is the number of deaths. This equation simply states that the frail population on day t comprises those frail individuals less the number of deaths from the previous day plus the newly frail. Hence on days following a large number of deaths, the very frail population at high risk of pollution-induced death is reduced and fewer deaths can result.

Given N_t frail persons on day t, we assume that each person independently experiences a daily hazard of death λ_t, where the log-odds of death depends linearly on the pollution value x_t. Finally, we assume that there is an effectively infinite total population, and that the number of persons, I_t, that enter the frail subpopulation at the start of day t follows a Poisson distribution with mean

$$\mu_I \left(\frac{I^*_{t-1}}{\mu_I} \right)^{\alpha},$$

where I^*_{t-1} is either the number of entrants from the previous day or a small positive constant if $I_{t-1} = 0$. This model for I_t allows the number of persons entering the frail state on a given day to depend on the number that entered on the previous day, introducing some positive autocorrelation (when $\alpha > 0$) that might reflect the influence of such events as influenza epidemics or stretches of bad weather (Zeger and Qaqish 1988). We assume that the long-term average of persons entering the frail subgroup $E[I_t]$ and the long-term average of the frail persons dying per day $E[D_t]$

are equal, so the population neither grows nor shrinks in the long run. Finally, to initiate the mortality D_t and frailty N_t time series, we assume that the number of persons at risk on the first day follows a Poisson distribution with mean equal to the long term average $E[N_t]$.

To demonstrate that mortality displacement induces associations between mortality and pollution only at short time scales, we have generated 3 long (N_t, D_t) series, which represent different degrees of mortality displacement. We chose the parameter values in Table 1 so that the mean number of deaths per day is 50, similar to the Philadelphia data, and the mean residence times (MRTs) in the frail state are 3, 30, or 300 days. The log-relative risk of mortality associated with particles was 0.05 for all scenarios. These data were generated in a simulated example using the TSP series from Philadelphia for 608 days during the period 1980 to 1994 for x_t, with 100 repetitions to create the simulated series of length 60,800.

Figure 1 displays the association, as measured by the squared correlation (coherence) between the number of deaths (square root transformed) $\sqrt{D_t}$ and particle measurement x_t as a function of time-scale, for the 3 mortality displacement scenarios. These plots were generated by cross-spectral analysis (Bloomfield 1976). Note that in all 3 situations the correlation becomes nonnegligible only at time scales that are less than about twice the MRT. Hence, we see that mortality displacement induces association between mortality and pollution only at shorter time scales.

In the next 2 subsections, we review our approach to estimating the mortality-pollution association separately at multiple time scales and then use that approach to produce estimates of pollution relative risks that are resistant to mortality displacement by ignoring the short-term information that can be influenced by mortality displacement.

Table 1. Parameter Values for Simulated Example

	Mean Residence Time		
Parameter	3 Days	30 Days	300 Days
$E[N_t]$	150	1,500	15,000
$E[I_t]$	50	50	50
$E[D_t]$	50	50	50
00	−0.69	−3.36	−5.70
01	0.05	0.05	0.05
a	0.80	0.80	0.80
C	1	1	1

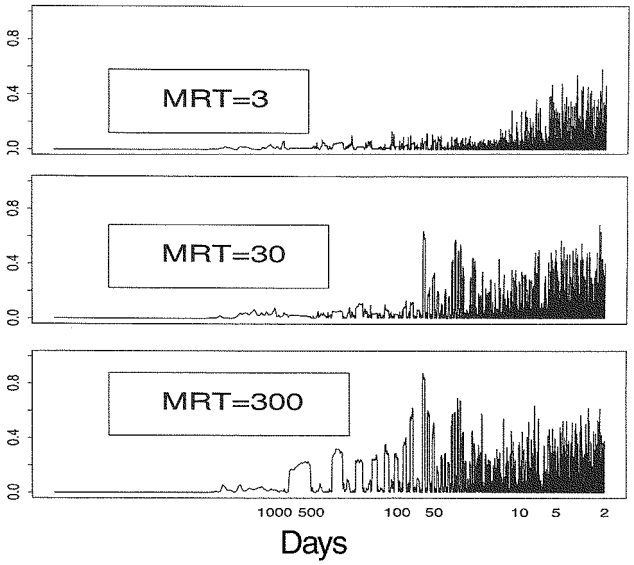

Figure 1. Estimated squared coherency between each pair of $\sqrt{D_t}$ and x_t.

FREQUENCY DOMAIN LOG-LINEAR REGRESSION

The main idea of FDLLR is to decompose both the air pollution series x_t and the mortality series (hereafter referred to as y_t) into distinct component series x_{kt} and y_{kt}, one pair for each of many distinct time scales k, and to calculate the association separately between x_{kt} and y_t for each time scale.

Figure 2 shows such a decomposition into 5 time scales—that is, roughly, year, season, month, week, day—for the TSP and mortality series from Philadelphia for the period 1974 to 1988. Note the top series comprise only the longest-term fluctuations, while the bottom series represents the shortest-term variations. The actual value of x_t (or y_t) on day t is obtained by summing the values of the 5-component series on that day. This type of decomposition can be obtained by smoothing with successively shorter running averages.

Frequency domain log-linear regression estimates a separate coefficient $\hat{\beta}_\kappa$ by regressing each y_t component on its corresponding component x_t giving a sequence of regression coefficients, for example, $\hat{\beta}_1$, $\hat{\beta}_2$, $\hat{\beta}_3$, $\hat{\beta}_4$, and $\hat{\beta}_5$ in this 5-component illustration.

The actual implementation of FDLLR uses a Fourier series decomposition of the x_t and y_t series and produces a smooth pollution-mortality log-relative risk function $\hat{\beta}_\kappa$ of the time scale or, equivalently, the frequency: k cycles in the total period of observations, rather than estimates at only 5 or some other small number of time scales. Kelsall and colleagues (1999) give a detailed description of FDLLR.

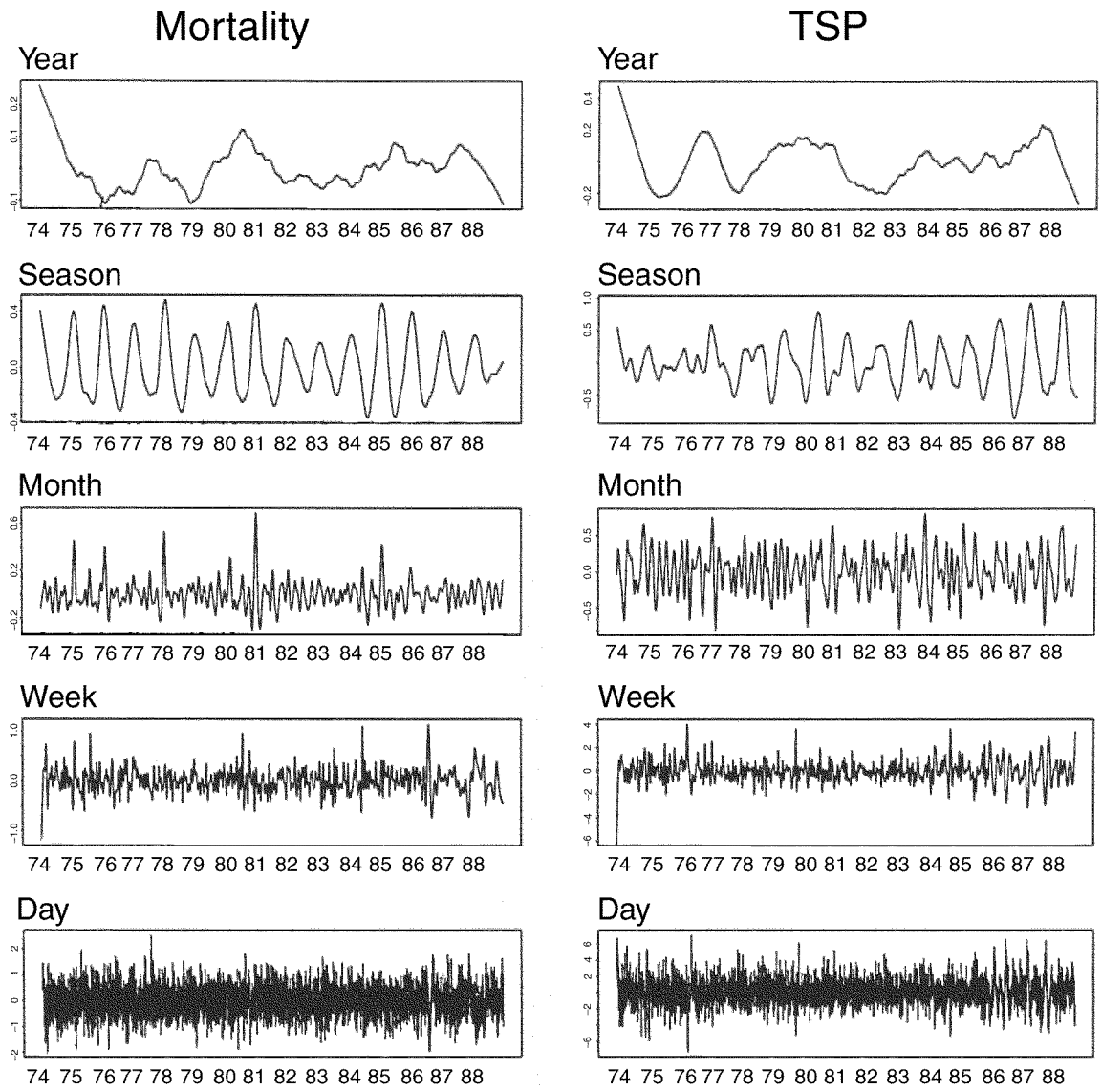

Figure 2. Decomposition into 5 component series for the mortality and TSP (μg/m^3) series (on squared root scale) from Philadelphia for the period 1974 to 1988. Each plot represents the residual time series with respect to the previous component.

MORTALITY DISPLACEMENT–RESISTANT ESTIMATES OF MORTALITY-POLLUTION RELATIVE RISK

As we demonstrated above, mortality displacement will affect only $\hat{\beta}_\kappa$ for large k, say $k > K$. In the simulation study we set the time scale k equal to twice the mean residence time in the frail state. Under the hypothesis that mortality displacement is the only cause of the pollution-mortality association, we would expect $\hat{\beta}_\kappa$ to be near 0 at low frequencies (when k is small) and to increase in absolute value toward higher frequencies (shorter time scales). A plot of $\hat{\beta}_\kappa$ versus k is therefore informative with regard to the mortality displacement hypothesis. We can also calculate a single mortality displacement–resistant estimator of β by taking an appropriately weighted average of the $\hat{\beta}_\kappa$ for $1 \le k \le K$, ignoring information at higher frequencies. The estimator is specifically defined as:

$$\hat{\beta}_\kappa = (\Sigma_{k \le K} w_k \bar{x}_k x_k)^{-1} (\Sigma_{k \le K} w_k \bar{x}_k z_k) \quad (1)$$

where x_k and z_k are the discrete Fourier transforms of x_t and of the linearized response z_t used in generalized linear

models (Zeger and Qaqish 1988), $-\boldsymbol{w}_k^{-1} = \mathbf{Var}(z_k)$, and $\bar{\boldsymbol{x}}_t$ is the complex conjugate of x.

Many pollution-mortality analyses use daily time-series data (Kelsall et al 1997). With series of n days, the possible values of k range from 1 to $n/2$ complete cycles in n days corresponding to frequencies: 1 cycle in n days to 1 cycle in 2 days. We seek a mortality displacement–resistant estimator that ignores the $\hat{\beta}_K$ corresponding to periods shorter than about twice the MRT in the frail state. We showed above that, under a simple model, the pollution-mortality associations are negligible at the remaining time scales when mortality displacement is the only source of this association.

When our mortality displacement–resistant estimator is applied to the simulated data with K corresponding to twice the mean residence time, the estimated pollution effect is close to 0 in all 3 cases.

Table 2 shows the fraction of the total frequencies corresponding to the shorter-time scales that must be ignored to protect against mortality displacement for various values of the MRT in the frail state given daily time-series data. A large fraction of the available information is affected by mortality displacement. Nevertheless, the remaining information at the longer frequencies can be used to estimate the pollution-mortality association without being biased by short-term mortality displacement, as illustrated for Philadelphia.

Table 2. Fraction of Frequencies Whose Corresponding Periods Are Shorter Than Twice Mean Residence Time (MRT) in Frail State

MRT (Days)	Frequencies to Ignore (%)
2	50
4	75
8	87.5
16	93.75

APPLICATION TO PHILADELPHIA DATA

To illustrate our method, we consider mortality and air pollution in Philadelphia for the years 1974 to 1988. The data set is the same as that used by Samet and colleagues (1997). In addressing the mortality displacement question, we implement our FDLLR while adjusting for temperature, dew point, and longer-term trends associated with factors such as changes in medical practice, demographics, and influenza epidemics. The adjustment is actually implemented using smoothing splines (Hastie and Tibshirani 1990) with 6 degrees of freedom (*df*) for temperature and dew point and with 90 *df* for time, respectively. The adjustment is similar to but not identical to those by Samet and colleagues (1997), who used 120 *df* for time and included day of the week as well.

Our goal is to calculate estimates of the pollution-mortality association in time-series data that sets aside, or ignores, information affected by mortality displacement. We first apply the FDLLR methods that estimate the pollutant-mortality association separately at each time scale.

On the left of Figure 3 we show the time-scale-specific estimates of the mortality relative risk associated with TSP by time scales. The horizontal axis is the time scale in days at which the association is measured. The solid and dotted lines are the estimated log-relative risks ±2 estimated standard errors, respectively, at each time scale. The plot is scaled to show the expected change in mortality corresponding to a change of 1 interquartile range (IQR) in TSP. On the right of Figure 3, we show the estimated mortality displacement–resistant effects of TSP on mortality as successively more of the shorter-term information is removed. The estimate at a particular time scale is calculated disregarding information in the left-hand panel at all shorter time scales.

Note that the pattern in the left panel is the opposite of the expectation under the mortality displacement hypotheses. The pollution relative rate is substantially different

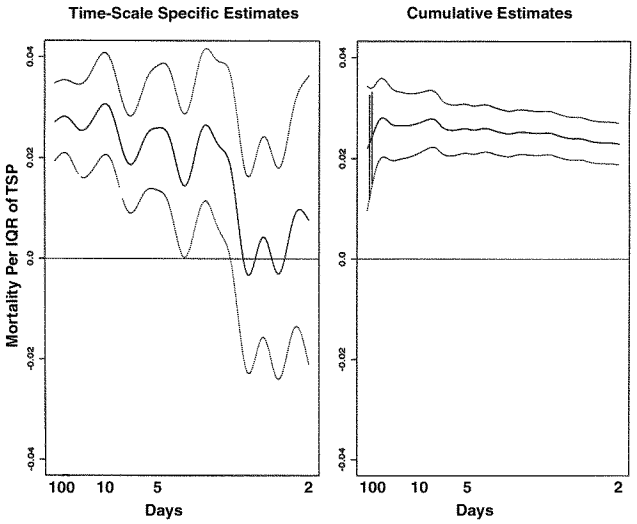

Figure 3. Time-scale (frequency) specific and cumulative estimates of the log-relative risk for mortality associated with current-day total suspended particles versus time scales in days estimated from daily time-series data from Philadelphia for 1974 to 1988. The solid and dotted lines are the estimates ±2 estimated standard errors. To the left of the cumulative estimates plot, mortality displacement–resistant estimators correspond to a mean residence time of 2 days and 4 days, respectively. The cumulative estimates at a certain time scale K (days/cycle) is calculated disregarding time scales shorter than K.

from 0 at low frequencies, and in fact, decreases rather than increases toward shorter-term frequencies. This pattern is opposite to what is expected under the mortality displacement–only hypothesis. The mortality displacement–resistant estimators corresponding to MRTs of 2 and 4 days are 0.022 (95% CI, 0.012 to 0.032) and 0.024 (95% CI, 0.015 to 0.033), respectively, indicating that the association between TSP and mortality reflects factors other than mortality displacement alone. Figure 3 reformats the information in Figure 2 into cumulative estimates that set aside varying amounts of the short-term information—moving from left to right, the estimate falls as increasing amounts of short-term data are used.

DISCUSSION

The method we present can be approximated by using filtering techniques to split the pollutant and mortality time series into components that have variations on distinct time scales, as was done in Figure 2. A log-linear regression of the component mortality on the corresponding component pollution series can be performed to obtain a separate relative-risk estimate for each component pair. Our method has the advantage of giving relative-risk estimates that are continuous functions of time scale rather than providing only a few discrete values. It also provides valid CIs that are not directly available from log-linear regression programs.

In the current implementation of our mortality displacement–resistant estimator, we fix the time lag between the pollution-exposure and mortality, rather than estimate it from the data. Hence, it is advisable to consider multiple, reasonable lags. For Philadelphia, the results were qualitatively the same when the TSP were lagged 0, 1, 2, or 3 days.

In gauging the public health significance of the evidence from the daily time-series studies, we find that the extent of mortality displacement has been a major point of controversy. With little evidence of significant life shortening, the findings from the time-series studies may not warrant a regulating response. In the recent standard-setting process for particulate matter, 2 cohort studies have figured prominently because their findings indicate longer-term effects (Dockery et al 1993; Pope et al 1995). Using a new analytic approach, our reassessment of the Philadelphia data indicates that the previously reported associations between air pollution indicators and mortality cannot be attributed solely to mortality displacement. This analytic approach should be extended to additional data sets to assess the consistency of our findings in Philadelphia.

REFERENCES

Bascom R, Bromberg PA, Costa DA, Devlin R, Dockery DW, Frampton MW, Lambert W, Samet JM, Speizer FE, Utell M. 1996a. Health effects of outdoor air pollution. Part 1. Am J Respir Crit Care Med 153:3–50.

Bascom R, Bromberg PA, Costa DA, Devlin R, Dockery DW, Frampton MW, Lambert W, Samet JM, Speizer FE, Utell M. 1996b. Health effects of outdoor air pollution. Part 2. Am J Respir Crit Care Med 153:477–498.

Beaver H. 1953. Interim Report (on London Fog Incident of 1952). Commission on Air Pollution: CMD9011. Her Majesty's Stationery Office, London, England.

Bloomfield P. 1976. Fourier Analysis of Time Series: An Introduction. John Wiley & Sons, New York NY.

Dockery DW, Pope CA III. 1994. Acute respiratory effects of particulate air pollution. Annu Rev Public Health 15:107–132.

Dockery DW, Pope CA III, Xu X, Spengler JD, Ware JH, Fay ME, Ferris BG Jr, Speizer FE. 1993. An association between air pollution and mortality in six US cities. N Engl J Med 329:1753–1759.

Hastie TJ, Tibshirani RJ, eds. 1990. Generalized Additive Models. Chapman & Hall, New York NY.

Kelsall JE, Samet JM, Zeger SL, Xu J. 1997. Air pollution and mortality in Philadelphia, 1974–1988. Am J Epidemiol 146:750–762.

Kelsall J, Zeger S, Samet J. 1999. Frequency domain log-linear models: Air pollution and mortality. Appl Stat 48 (Pt 3):331–344.

Lipfert FW, Wyzga RE. 1995. Air pollution and mortality: issues and uncertainties. J Air Waste Manage Assoc 45: 949–966.

Pope CA III, Dockery DW, Schwartz J. 1995. Review of epidemiological evidence of health effects of particulate air pollution. Inhalation Toxicol 7:1–18.

Samet JM, Zeger SL, Berhane K. 1995. The association of mortality and particulate air pollution. In: Particulate Air Pollution and Daily Mortality: Replication and Validation of Selected Studies, The Phase I.A Report of the Particle Epidemiology Evaluation Project. Health Effects Institute, Cambridge MA.

Samet JM, Zeger DL, Kelsall JE, Xu J, Kalkstein LS. 1997. Air pollution, weather, and mortality in Philadelphia

1973–1988. In: Particulate Air Pollution and Daily Mortality: Analyses of the Effects of Weather and Multiple Air Pollutants, The Phase I.B Report of the Particle Epidemiology Evaluation Project. Health Effects Institute, Cambridge MA.

Schimmel H, Murawski TJ. 1976. Proceedings: The relation of air pollution to mortality. J Occup Med 18:316–333.

Schwartz J, Dockery DW. 1992. Increased mortality in Philadelphia associated with daily air pollution concentrations. Am Rev Respir Dis 145:600–604.

Smith L, Davis J, Sacks J, Speckman P, Styer P. 1997. Assessing the human health risk of atmospheric particles. In: Statistics and the Environment. Proceedings from the 1997 Joint Statistical Meetings, Anaheim CA. American Statistical Association, Alexandria VA.

US Environmental Protection Agency. 1996. Air Quality Criteria for Particulate Matter. EPA/600/P-95/001 af. Government Printing Office, Washington DC.

US Environmental Protection Agency. 1996. Review of the National Air Quality Standards for Particulate Matter: Policy Assessment of Scientific and Technical Information. OAQPS Staff Paper. EPA-452\R-96-013. Office of Air Quality Planning and Standards, Research Triangle Park NC.

Zeger SL, Qaqish B. 1988. Markov regression models for time series: A quasi-likelihood approach. Biometrics 44:1019–1031.

Section 4: Mortality Displacement and Long-term Exposure Effects Related to Air Pollution and Mortality

Joel Schwartz

ABSTRACT

Although time-series analyses have demonstrated that airborne particles are associated with early death, these analyses have not clarified by how much the deaths are advanced. If all of the pollution-related deaths were advanced by only a few days, one would expect little association between weekly averages of air pollution and daily deaths. I used the STL (seasonal and trend decomposition using LOESS [locally weighted smoother])* algorithm to resolve air pollution, daily deaths, and weather from Boston into 3 time series: one reflecting seasonal and longer fluctuations, one reflecting short-term fluctuations, and one reflecting intermediate patterns. By varying the cutoff between short term and intermediate term, it was possible to look at mortality displacement on different time scales. For chronic obstructive pulmonary disease (COPD), there was evidence that most of the mortality was displaced by only a few months. For pneumonia, heart attacks, and all causes of mortality, the size of the effect increased at longer time scales. The percent increase in all deaths associated with a 10 $\mu g/m^3$ increase in particulate matter less than 2.5 μm in aerodynamic diameter ($PM_{2.5}$) increased from 2.1% (1.5–4.3) to 3.75% (3.2–4.3) as the focus moved from daily to monthly patterns. This is consistent with the larger effect seen in prospective cohort studies, rather than mortality displacement playing a major role.

* A list of abbreviations and other terms appears on page 13.

The National Morbidity, Mortality and Air Pollution Study: Methods and Methodologic Issues, Part I of Health Effects Institute Research Report 94, includes an Investigators' Report, a Preface, a Commentary by the Health Review Committee, and an HEI Statement about the research project. Correspondence concerning this section may be addressed to Dr Joel Schwartz, Department of Environmental Health, Harvard School of Public Health, 665 Huntington Avenue, I-1414, Boston MA 02115-6021.

Although this document was produced with partial funding by the United States Environmental Protection Agency under Assistance Award R824835 to the Health Effects Institute, it has not been subjected to the Agency's peer and administrative review and therefore may not necessarily reflect the views of the Agency, and no official endorsement by it should be inferred. The contents of this document also have not been reviewed by private party institutions, including those that support the Health Effects Institute; therefore, it may not reflect the views or policies of these parties, and no endorsement by them should be inferred.

INTRODUCTION

A large body of literature has shown associations between particulate air pollution and daily mortality and morbidity (Bates and Szito 1987; Pope 1989; Fairley 1990; Schwartz and Marcus 1990; Pope et al 1991, 1992, 1995a; Schwartz 1991, 1993a, 1994a,b,c,d, 1995b, 1996, 1997; Schwartz and Dockery 1992; Schwartz et al 1993, 1996; Spix et al 1993; Sunyer et al 1993; Burnett et al 1994; Thurston et al 1994; Schwartz and Morris 1995; Verhoeff et al 1996; Anderson et al 1997; Katsouyanni et al 1997; Zeger et al 1999). These associations have been shown in locations or seasons where O_3 and SO_2 concentrations were essentially nonexistent, making confounding by these other pollutants implausible (Schwartz 1995a). As a result, national and international bodies (Expert Panel on Air Quality Standard Particles 1995; US Environmental Protection Agency [EPA] 1997; World Health Organization 1997) have concluded that these associations should be treated as causal and have recommended implementing tighter standards. Although the association between PM and daily mortality is generally accepted, considerable controversy exists about the extent to which deaths are advanced by higher air pollution levels. Some have argued that it is inappropriate to use the regression coefficients from these studies to estimate the attributable risk of air pollution because it is not clear how many of the deaths occur only a few days early in individuals who are already dying (McMichael et al 1998). This is usually referred to as *harvesting* or *mortality displacement*. If all of the deaths associated with particulate air pollution were being displaced by only a few days, this would obviously have implications for the extent of public health concern that should be given to the associations.

On the other hand, the existing studies have examined the association of air pollution exposure on mortality and morbidity only on the same or past few days. Exposure of animals to combustion particles indicates that they produce inflammatory damage in the lung at least partially by the generation of oxidants (Gilmour et al 1996; Li et al 1996; Pritchard et al 1996; Costa and Dreher 1997). This suggests that exposure over time intervals of weeks may have some additional cumulative effect that is not captured in the

current short-term regression analyses. Further, studies showing that air pollution is associated with increased severity of illness (Bates and Szito 1987; Pope 1989; Pope et al 1991; Schwartz et al 1993; Sunyer et al 1993; Burnett et al 1994; Schwartz 1994a,b,c, 1995b, 1996, 1997; Thurston et al 1994; Schwartz and Morris 1995; Anderson et al 1997) suggest that it can increase the pool of persons at high risk of dying (by moving people from moderate to severate illness) as well as deplete it. The net effect on the size of the susceptible pool is not clear a priori. Finally, prospective cohort studies of particulate air pollution and daily deaths (Dockery et al 1993; Pope et al 1995b) have reported substantially larger effects of long-term exposure to, for example, 10 $\mu g/m^3$ of fine particles than are indicated by the daily time-series studies. Those authors have suggested that the difference may represent an effect of chronic exposure. Section 4 seeks to examine how the association between particulate air pollution and mortality and morbidity varies as the time scale of the exposure varies.

DATA

Mortality data came from Boston for the years 1979 to 1986 as described previously (Schwartz et al 1996). Briefly, between 1979 and 1986, dichotomous virtual impactor samplers were placed at a central, residential monitoring site in the Boston metropolitan area as part of the Harvard Six Cities Study. Separate samples of fine particles ($PM_{2.5}$) and coarse mass (CM) were collected. This analysis is restricted to the fine particle data. Daily deaths were extracted from annual detail mortality tapes from the National Center for Health Statistics (NCHS) for the same time period, for Norfolk, Suffolk, and Middlesex Counties, which are the metropolitan counties proximate to the monitor. Deaths due to accidents and other external causes (ICD-9 800–999) were excluded. Separate counts were computed for deaths from ischemic heart disease (IHD) (ICD-9 410–414), congestive heart failure (ICD-9 428), pneumonia (ICD-9 480–486), and COPD (ICD-9 490–496). Meteorologic data were obtained from the National Center for Atmospheric Research. The hourly measures were collapsed over 24-hour periods to provide a mean value for temperature and dew point. The initial paper reported an association between daily deaths and the average of air pollution on the same and previous days. The association was seen after control for weather (using nonparametric smooth functions of temperature and humidity) and for season (using a nonparametric smooth function of time with a smoothing window of about 125 days). It also examined how the association differed among different particle measures, which is not the topic of this analysis.

METHODS

If air pollution were advancing deaths by only a few days, then we would expect an increase in daily deaths due to air pollution to be followed shortly by a decline. If we averaged over a week, the 2 effects would cancel each other out (or partially cancel each other out if some of the deaths are brought forward by a longer period). Put another way, the association between air pollution and daily deaths would be concentrated in high-frequency fluctuations, those with periods of only a few days. A multiday average of daily deaths would no longer be associated with air pollution, since the air pollution effect and the rebound from it would have been smoothed over by the averaging. If we can separate the correlation between air pollution and daily deaths into characteristic frequency ranges, the existence of an association at lower frequencies would demonstrate that all of the air pollution–associated deaths are not being advanced by only a few days. This is the basis of the analysis. In contrast, if there are cumulative effects of exposure that are not captured in the daily regression analyses, or if air pollution increased the pool of susceptible individuals, the association between longer period fluctuations in air pollution and daily deaths would be stronger than in the original analysis. By examining this association in different frequency ranges, one can examine the existence of mortality displacement and effects of longer-term exposure on a range of time scales from a few days to a month or two. Examination of much longer time scales is difficult because of the need to control for season.

Cleveland's STL algorithm (Cleveland et al 1990) was used to separate the time series of daily deaths, air pollution, and weather into long wavelength components (representing time trends and seasonal fluctuations), midscale components, and residual very short time scale components. This analysis uses the midscale components of each time series to assess the association between air pollution and mortality on that scale, having removed the potentially confounding effect of season (long scale) and the component susceptible to short-term mortality displacement (short scale). The STL uses LOESS smoothing to separate the series into these components.

All analyses used the same cutoff for the long wavelength component. A LOESS smooth with a window of 120 days was used to fit and remove the seasonal and long-term time trends. The LOESS smooth uses a weighted moving regression within the 120-day window to estimate the seasonal component of variation for each time series (ie, deaths, $PM_{2.5}$, temperature, and dew point). The weights decrease to 0 at the ends of the window as the

cube of the fraction of the distance from the center to the end (see Appendix A), and are near 1 for only about the central 40% of the window. Hence the effectiveness of a 120-day window in a LOESS smooth at removing long wavelength patterns is similar to a simple unweighted moving average of about 60 days. The LOESS smooth is preferred because the weighted smoothing produces less distortion in the high-frequency components. Smaller window sizes (eg, 90 days) induce short-term serial correlation in the data that is not present in the original series.

Because the goal of the analysis was to examine the association in different frequency ranges, several different midscale components were examined separately. These midscale smoothing windows were 15, 30, 45, and 60 days. For each midscale window, the analysis was repeated, removing the seasonal and the short-term patterns from the data. Regression analysis was then performed among the midscale variations in deaths, pollution, and weather, for each of the 4 choices of midscale variation.

To maintain comparability with the original study, the same generalized additive model and choice of lags was used in these analyses. A log-linear regression was fit relating the logarithm of the filtered daily deaths (with the mean added back) to LOESS smooth functions of temperature, dew point, and a linear $PM_{2.5}$ term, for each of the different midscale frequency ranges. The smooth functions of temperature and dew point used approximately 5 degrees of freedom (*df*) each (see Appendix A for more details). The results for each of these windows were compared with results in the original regressions. This allows a comparison of the size of the effects as we sequentially exclude longer and longer term mortality displacement from the analysis.

Table 1. Mortality and Environment in Boston, 1979 to 1986

	Mean ± SD
Mortality	
All causes	60.0 ± 9.6
Pneumonia	2.7 ± 1.9
COPD	1.4 ± 1.4
IHD	17.9 ± 5.3
Environment	
Temperature (°C)	10.6 ± 9.6
$PM_{2.5}$ (μg/m^3)	15.6 ± 9.2
Dew point (°C)	4 ± 10.7

RESULTS

Table 1 shows the mean and standard deviations of the environmental and mortality data in Boston. Air pollution levels were low to moderate. Total deaths averaged 60. Figure 1 shows the 120-day LOESS smooth, which appears to capture seasonal variation and some shorter-term structure in the mortality data. Figure 2 shows the residuals after removing this pattern from the mortality data, confirming that no seasonal pattern was left in the data. In fact, the partial autocorrelation function was reduced to white noise by a 150-day LOESS smooth, and the 120-day window was used to be conservative. Figure 3 is an example of what data filtered in such a manner look like. It shows the residual daily deaths from IHD, after removing both seasonal and short-term fluctuations, over time. The mean is 0 in the figure, but the mean was added back in to

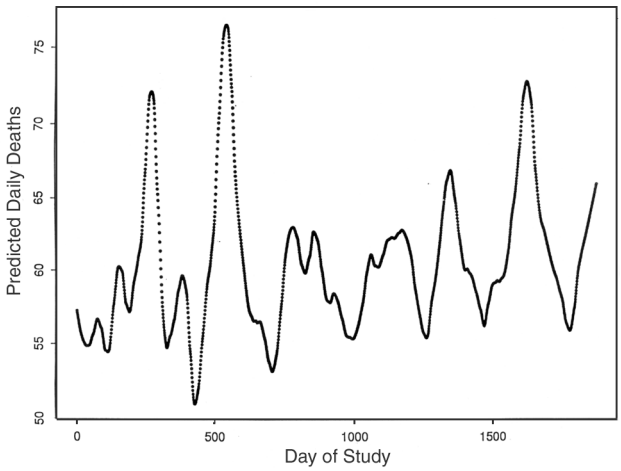

Figure 1. Results of the 120-day window LOESS smooth function of daily deaths in Boston over time for the period 1979 to 1986. This smooth was removed from the data to control for season and trend in all of the analyses.

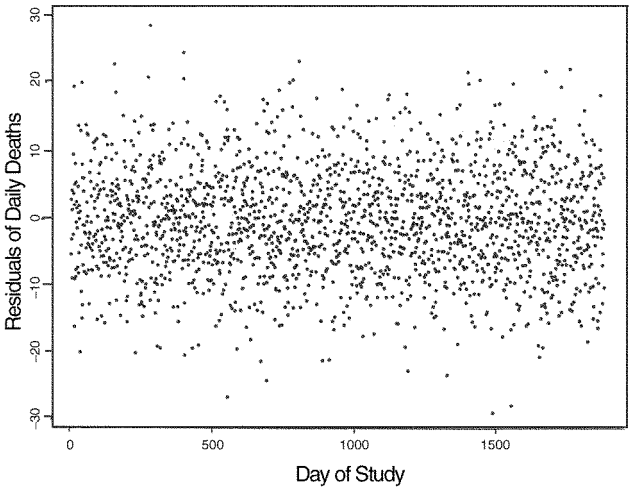

Figure 2. Residual mortality after seasonal control. Shows the residuals of the seasonal smooth in Figure 1. No seasonality is apparent in the residuals.

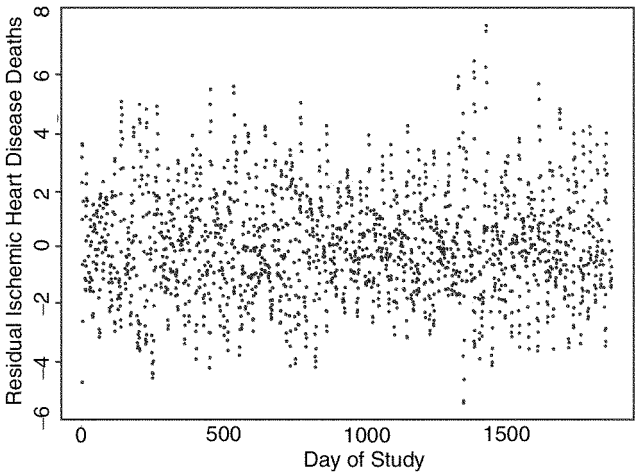

Figure 3. Plot of the residuals of ischemic heart disease (IHD) deaths versus day of study, after removing seasonality using a 120-day LOESS filter and short-term fluctuations using a 15-day LOESS filter.

perform the log-linear regressions. Figure 4 shows the estimated increase in death from COPD (and 95% confidence interval [CI]) associated with a 10 μg/m^3 increase in $PM_{2.5}$ (from Schwartz et al 1996) and with each of the 4 filters. Figures 5 and 6 show the same thing for pneumonia and IHD deaths. These 3 cause-specific plots illustrate different patterns of association.

The results for COPD show a pattern similar to what has been hypothesized about mortality displacement. The effect size first increases when a 15-day smoothing window is used, and then decreases to 0 by the 60-day smoothing window. This suggests that the deaths from COPD are being advanced by only a few weeks to a few months.

Pneumonia, in contrast, shows some sign of short-term mortality displacement, with a lower effect size with a 15-day smoothing window, but then the effect size grows to more than twice the original estimate by the time a 60-day smoothing window is reached. This pattern is not consistent with most of the deaths being advanced by a few days to few months.

For IHD deaths, the effect size is unchanged when using the 15-day window. For larger averaging windows, the effect size increases monotonically. For the 60-day window, which focuses the association on correlations with a 30-day to 100-day time scale, the effect of air pollution is almost twice as great as in the original regression.

The results for all-cause mortality most strongly resemble those for IHD. The 15-day smoothing window results in little change, but the effect size then increases steadily with increasing averaging times. This is indicated in Figure 7.

DISCUSSION

The findings in Boston provide some evidence to justify both short-term mortality displacement effects as well as larger effects when short-term mortality displacement is excluded. These latter effects may be the effects of longer-

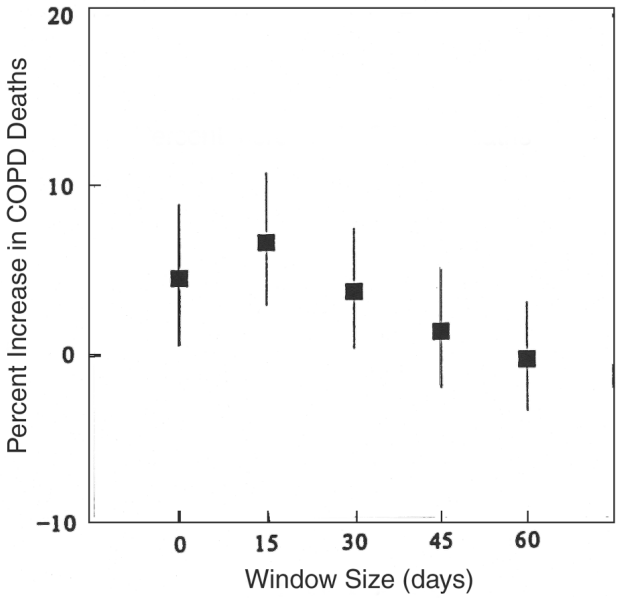

Figure 4. Effect on COPD mortality. The estimated effect of a 10 μg/m^3 increase in $PM_{2.5}$ concentration on daily mortality from COPD in Boston (Schwartz et al 1996) and for the 4 analyses in this study, using windows of 15, 30, 45, and 60 days.

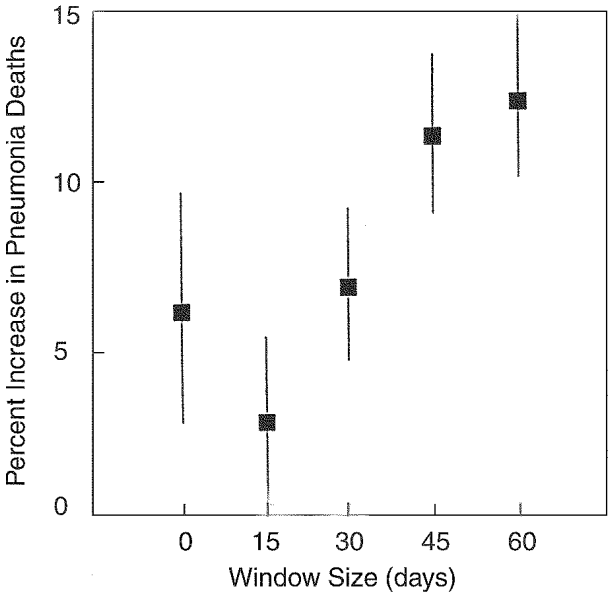

Figure 5. Effect on pneumonia mortality. The estimated effect of a 10 μg/m^3 increase in $PM_{2.5}$ concentration on daily mortality from pneumonia in Boston (Schwartz et al 1996) and for the 4 analyses in this study, using windows of 15, 30, 45, and 60 days.

term average pollution concentrations, or they may reflect increased recruitment into the susceptible pool of individuals caused by air pollution. When comparing with the original regression results, we should recall that once seasonal patterns are removed, the greatest variation in the air pollution data occurs on time scales of less than a week. Hence the regression is presumably dominated by results in that time scale. As we move to regressing the filtered time series, we switch the dominant time scale first to variations over a few weeks, and eventually to variations over a few months. By that time, for COPD, the effect of air pollution has disappeared. This suggests that the COPD deaths are being brought forward usually by a few months. It should be noted that these results apply to deaths where COPD is listed as the underlying cause. Chronic obstructive pulmonary disease may be a contributing cause to deaths with other underlying causes listed, and the pattern may differ in that case. Rats with COPD had excessive mortality when exposed to 200 to 300 $\mu g/m^3$ of particles in exposure chambers in a recent study (Godleski et al 1996); however, they died in their sleep without signs of respiratory distress, and the deaths may have been due to cardiovascular effects.

In contrast, the pneumonia results suggest that there may be some deaths brought forward by a few days, which produces the diminished effect on a time scale of a few weeks, but that this effect is overwhelmed by the larger effect sizes when all longer-term filters are applied. Over a time scale of 1 or 2 months, the effect of air pollution seems substantially larger on pneumonia deaths than originally reported. This is not surprising. People with pneumonia rarely linger on the edge of death for months. If the pneumonia is potentially life threatening, it usually remains so for a limited period, followed by recovery or death. If patients recover from pneumonia, they are probably safe until the next episode of pneumonia, which is likely to be a year or more in the future. Pneumonia hospital admissions in Chicago in 1992 confirm this. Of persons aged 65 and older who were admitted to Chicago-area hospitals in January and February, only 8% had a readmission in the next 6 months. Hence a pattern of some, but not most, deaths from pneumonia being brought forward by a few days makes sense. The possibility that longer-term exposures to particulate air pollution may exacerbate pneumonia deaths is also plausible, since particulate exposure is associated with inflammatory processes. Moreover, the association between particulate air pollution and pneumonia hospital admissions (Cleveland et al 1990; Godleski et al 1996; Zelikoff et al 1997) suggests the pool of persons at risk of dying from pneumonia may be increased by particulate air pollution, not decreased. Animal studies have shown that exposure to combustion particles exacerbates pneumonia in rats (Zelikoff et al 1997) and influenza infections in mice (Clarke et al 1997), lending further credence to this association. It is still possible that the deaths of some of these individuals, particularly at older ages, are

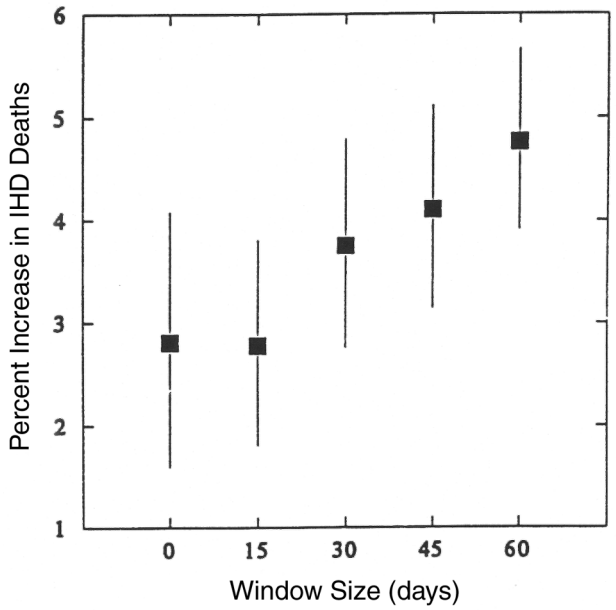

Figure 6. Effect of PM concentration on IHD mortality. The estimated effect of a 10 $\mu g/m^3$ increase in $PM_{2.5}$ concentration on daily mortality from ischemic heart disease in Boston (from Schwartz et al 1996) and for the 4 analyses in this study, using windows of 15, 30, 45, and 60 days.

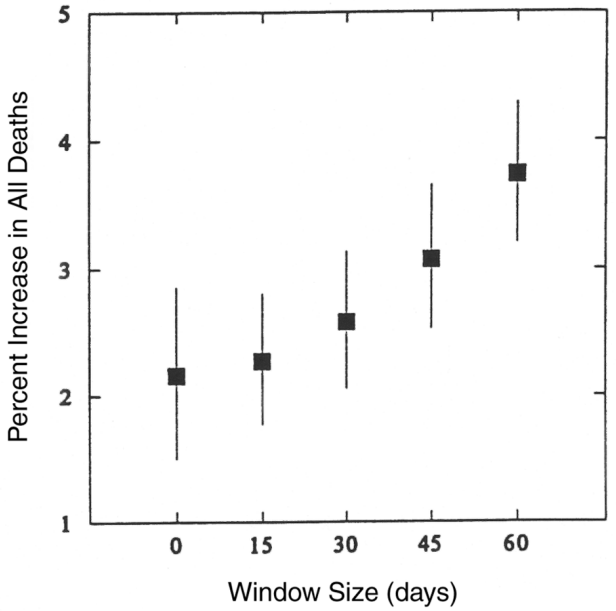

Figure 7. Effect of PM concentration on all-cause mortality. The estimated effect of a 10 $\mu g/m^3$ increase in $PM_{2.5}$ concentration on all cause mortality in Boston (from Schwartz et al 1996) and for the 4 analyses in this study, using windows of 15, 30, 45, and 60 days.

being brought forward by only several months, which is still a modest amount. However, the natural history of pneumonia suggests that most of the people who recover from pneumonia will not contract another case until the next pneumonia season.

For IHD and for mortality from all causes, excluding short-term changes definitely leads to an increase in the estimated effect of air pollution. If one thinks of heart attacks as Poisson events in vulnerable populations, then it is not surprising that if an event is avoided on a given day, the expected displacement of mortality will be greater than months. The effect of air pollution might be primarily to exacerbate a myocardial infarction brought on by other stimuli. Here too, the analysis cannot exclude the possibility that the deaths are being brought forward by, for example, only 3 months. Since the 5-year survival rate for people who survive the first 48 hours of a heart attack is quite high, however, this is unlikely to be the case for most of the avoided early deaths. Again, of elderly persons in Chicago admitted to hospital for myocardial infarctions in January and February of 1992, only 6% had a second admission for a myocardial infarction in the next 6 months. Pope and coworkers (1995b) examined the relationship between fine particle exposure on a scale of years and deaths in a prospective follow-up study involving most of the urban areas in the United States. They reported that a 10 $\mu g/m^3$ increase in $PM_{2.5}$ concentration was associated with a 6.6% increase in all-cause mortality. They attributed the difference between that effect estimate and results such as the 2.1% estimate seen in the original time-series study from Boston (Schwartz et al 1996) as suggesting a greater effect of long-term exposure, possibly due to the development of chronic disease.

Other studies have indicated that particulate exposure is a risk factor for the development of COPD, for example (Euler et al 1987; Schwartz 1993b). Some have argued that the higher slope observed by Pope and colleagues (1995b) reflects the higher exposures that existed 20 years earlier in their study locations (Lipfert 1997). This analysis indicates that moving from a time scale of days to months captures about half the difference between the daily time series and the prospective cohort study. This suggests that most of the increase in slope occurs over relatively short time scales and does not take 20 years of exposure to develop. Of course, it is also possible that the higher slope in the cohort studies results from uncontrolled confounding.

There is a developing literature on potential mechanisms by which particulate air pollution might affect the cardiovascular system. Exposure of dogs to 100 to 200 $\mu g/m^3$ of fine particles in an exposure chamber for 6 hours per day for 3 days resulted in electrocardiogram changes that are risk factors for arrhythmia (Godleski et al 1997). These changes were enhanced in the presence of preexisting angina (Godleski et al 1997). Instillation of 250 µg of combustion particles into the lungs of rats produced arrhythmias and deaths in another recent study (Watkinson et al 1998). In humans, particulate air pollution has been associated with increases in plasma viscosity (Peters et al 1997), increased risk of elevated heart rate (Pope et al 1999), and changes in heart-rate variability (Shy et al 1997).

It would be of interest to examine whether the deaths associated with particulate air pollution continue to increase as the averaging time increases further. In particular, how long does it take to reach the levels seen in the prospective cohort studies? Seasonality, however, impedes examining this further using time-series data. The use of a filter to remove seasonality prevents us from examining longer-term averaging periods. One alternative would be to use a larger window to control for season, but this would increase the risk of confounding by inadequately controlling for season, and hence represents an inherent limitation of such time-series analyses. Another limitation of the study is the choice of lags for the exposure variables. The original study chose a priori to use the mean of the pollution on the same and previous day in all 6 locations studied. I have repeated that choice to maintain comparability. The original paper also used weather variables on the same day for each city studied. Further examination in Boston could reveal a better fit from a different weather model. Again the same model was used to maintain comparability. In this way differences in effect size estimates can be uniquely attributable to discarding the higher frequency variations in the data. Kelsall and coworkers (1997) have reported that the association between airborne particles and daily deaths in Philadelphia was insenstive to variations in control for season and weather.

Overall, the results of this study suggest that the time-series studies that have been published underestimate rather than overestimate the number of early deaths that are associated with air pollution and that are brought forward by nontrivial amounts of time.

ACKNOWLEDGMENTS

The idea for this work arose in part in discussions with Dr Scott Zeger about work he was doing using frequency domain regression to address mortality displacement (Zeger et al 1999).

REFERENCES

Anderson HR, Spix C, Medina S, Schouten J, Castellsague J, Rossi G, Zmirou D, Touloumi G, Wojtyniak B, Ponka A, Bacharova L, Schwartz J, Katsouyanni K. 1997. Air pollution and daily admissions for chronic obstructive pulmonary disease in 6 European cities: Results from the APHEA project. Eur Respir J 10:1064–1071.

Bates DV, Szito R. 1987. Hospital admissions and air pollutants in Southern Ontario: The acid summer haze effect. Environ Res 43:317–331.

Burnett RT, Dales RE, Raizenne ME, Krewski D, Summers PW, Roberts GR, Radd-Young M, Dann T, Brook J. 1994. Effects of low ambient levels of ozone and sulfates on the frequency of respiratory admissions to Ontario hospitals. Environ Res 65:172–194.

Clarke RW, Hemenway DR, Frank R, Kleeberger SR, Longphre MV, Jakab GJ. 1997. Particle-associated sulfate exposure enhances murine influenza mortality. Am J Respir Crit Care Med 155(Pt 2):A245.

Cleveland RB, Cleveland WS, McRae JE, et al. 1990. Seasonal-trend decomposition procedure based on LOESS. J Off Stat 6:3–73.

Costa D, Dreher K. 1997. Bioavailable transition metals in particulate matter mediate cardiopulmonary injury in health and compromised animal models. Environ Health Perspect (Suppl 5) 105:1053–1060.

Dockery DW, Pope CA III, Xu X, Spengler JD, Ware JH, Fay ME, Ferris BG Jr, Speizer FE. 1993. An association between air pollution and mortality in six US cities. N Engl J Med 329:1753–1759.

Euler GL, Abbey DE, Magie AR, Hodgkin JE. 1987. Chronic obstructive pulmonary disease symptom effects of long-term cumulative exposure to ambient levels of total suspended particulates and sulfur dioxide in California Seventh-Day Adventist residents. Arch Environ Health 42:213–222.

Expert Panel on Air Quality Standard Particles, eds. 1995. A Recommendation for a United Kingdom Air Quality Standard for Particles. Her Majesty's Stationery Office, London, England.

Fairley D. 1990. The relationship of daily mortality to suspended particulates in Santa Clara County, 1980–1986. Environ Health Perspect 89:159–168.

Gilmour PS, Brown DM, Lindsay TG, Beswich PH, MacNee W, Donaldson K. 1996. Adverse health effects of PM_{10} particles: Involvement of iron in generation of hydroxy radical. Occup Environ Med 53:817–822.

Godleski JJ, Lovett EG, Sioutas C, Killingsworth CR, Krishnamurthi GG, Hatch V, Wolfson M, Ferguson ST, Koutrakis P, Verrier RL. 1997. Impact of inhaled concentrated ambient air particles on canine electrocardiographic patterns (abstract). In: The Thirteenth Health Effects Institute Annual Conference Program. Health Effects Institute, Cambridge MA.

Godleski JJ, Sioutas C, Katler M, Koutrakis P. 1996. Death from inhalation of concentrated air particles in animal models of pulmonary disease. Am J Respir Crit Care Med 153:A15.

Katsouyanni K, Touloumi G, Spix C, Schwartz J, Balducci F, Medina S, Rossi G, Wojtyniak B, Sunyer J, Bacharova L, Schouten JP, Ponka A, Anderson HR. 1997. Short-term effects of ambient sulphur dioxide and particulate matter on mortality in 12 European cities: Results from the APHEA project. Br Med J 314:1658–1663.

Kelsall JE, Samet JM, Zeger SL, Xu J. 1997. Air pollution and mortality in Philadelphia, 1974–1988. Am J Epidemiol 146:750–762.

Li XY, Gilmour PS, Donaldson K, MacNee W. 1996. Free radical activity and pro-inflammatory effects of particulate air pollution (PM_{10}) in vivo and in vitro. Thorax 51:1216–1222.

Lipfert F. Clean air skepticism. 1997. Science 278:19–20.

McMichael AJ, Anderson HR, Brunekreef B, Cohen AJ. 1998. Inappropriate use of daily mortality analyses to estimate longer-term effects of air pollution. Int J Epidemiol 27:450–453.

Peters A, Doring A, Wichmann HE, Koenig W. 1997. Increased plasma viscosity during an air pollution episode: A link to mortality? Lancet 349:1582–1587.

Pope CA III, Dockery DW, Kanner RE, Villegas GM, Schwartz J. 1999. Oxygen saturation, pulse rate, and particulate air pollution: A daily time-series panel study. Am J Respir Crit Care Med 159:365–372.

Pope CA III. 1989. Respiratory disease associated with community air pollution and a steel mill, Utah Valley. Am J Public Health 79:623–628.

Pope CA III, Dockery DW, Schwartz J. 1995a. Review of epidemiological evidence of health effects of particulate air pollution. Inhalation Toxicol 7:1–18.

Pope CA III, Thun MJ, Namboodiri MM, Dockery DW, Evans JS, Speizer FE, Heath CW Jr. 1995b. Particulate air

pollution as a predictor of mortality in a prospective study of US adults. Am J Respir Crit Care Med 151:669–674.

Pope CA III, Dockery DW, Spengler JD, Raizenne M. 1991. Respiratory health and PM_{10} pollution: A daily time-series analysis. Am Rev Respir Dis 144:668–674.

Pope CA III, Schwartz J, Ranson MR. 1992. Daily mortality and PM_{10} pollution in Utah Valley. Arch Environ Health 47:211–217.

Pritchard RJ, Ghio AJ, Lehmann JR, Winsett DW, Tepper JS, Park P, Gilmour MI, Dreher KL, Costa DL. 1996. Oxidant generation and lung injury after particulate air pollutant exposure incease with the concentrations of associated metals. Inhalation Toxicol 8:457–477.

Schwartz J. 1991. Particulate air pollution and daily mortality in Detroit. Environ Res 56:204–213.

Schwartz J. 1993a. Air pollution and daily mortality in Birmingham, Alabama. Am J Epidemiol 137:1136–1147.

Schwartz J. 1993b. Particulate air pollution and chronic respiratory disease. Environ Res 62:7–13.

Schwartz J. 1994a. Air pollution and hospital admissions for the elderly in Minneapolis. Arch Environ Health 49:366–374.

Schwartz J. 1994b. Air pollution and hospital admissions for the elderly in Birmingham, Alabama. Am J Epidemiol 139:589–598.

Schwartz J. 1994c. Air pollution and hospital admissions for the elderly in Detroit, Michigan. Am J Respir Crit Care Med 150:648–655.

Schwartz J. 1994d. What are people dying of on high air pollution days? Environ Res 64:26–35.

Schwartz J. 1995a. Health effects of air pollution from traffic: Ozone and particulate matter. In: Health at the Crossroads (McMichaels A, Fletcher T, eds). John Wiley & Sons, London, England.

Schwartz J. 1995b. Short-term fluctuations in air pollution and hospital admissions of the elderly for respiratory disease. Thorax 50:531–538.

Schwartz J. 1996. Air pollution and hospital admissions for respiratory disease. Epidemiology 7:20–28.

Schwartz J. 1997. Air pollution and hospital admissions for cardiovascular disease in Tucson. Epidemiology 8:371–377.

Schwartz J, Dockery DW. 1992. Increased mortality in Philadelphia associated with daily air pollution concentrations. Am Rev Respir Dis 145:600–604.

Schwartz J, Dockery DW, Neas LM. 1996. Is daily mortality associated specifically with fine particles? J Air Waste Manage Assoc 46:927–939.

Schwartz J, Marcus A. 1990. Mortality and air pollution in London: A time-series analysis. Am J Epidemiol 131:185–194.

Schwartz J, Morris R. 1995. Air pollution and hospital admissions for cardiovascular disease in Detroit, Michigan. Am J Epidemiol 142:23–35.

Schwartz J, Slater D, Larson TV, Pierson WE, Koenig JQ. 1993. Particulate air pollution and hospital emergency room visits for asthma in Seattle. Am Rev Respir Dis 147:826–831.

Shy CM, Creason J, Williams R, Liao D, Hazucha M, Devlin R, Zweidinger R. 1997. Physiological response in the elderly exposed to ambient particulate air pollution: A pilot study (abstract). In: The Thirteenth Health Effects Institute Annual Conference Program. Health Effects Institute, Cambridge MA.

Spix C, Heinrich J, Dockery DW, Schwartz J, Volksch G, Schwinkowski K, Collen C, Wichmann HE. 1993. Air pollution and daily mortality in Erfurt, East Germany, 1980–1989. Environ Health Perspect 101:518–526.

Sunyer J, Saez M, Murillo C, Castellsague J, Martinez F, Anto JM. 1993. Air pollution and emergency room admissions for chronic obstructive pulmonary disease. Am J Epidemiol 137:701–705.

Thurston GD, Ito K, Hayes CG, Bates DV, Lippmann M. 1994. Respiratory hospital admissions and summertime haze air pollution in Toronto, Ontario: Consideration of the role of acid aerosols. Environ Res 65:271–290.

US Environmental Protection Agency. 1997. National Ambient Air Quality Standard for Particulate Matter, Part KK. Fed Regist 62:138.

Verhoeff AP, Hoek G, Schwartz J, Van Wijnen JH. 1996. Air pollution and daily mortality in Amsterdam. Epidemiology 7:225–230.

Watkinson WP, Campen MJ, Costa DL. 1998. Cardiac arrhythmia induction after exposure to residual oily fly ash particles in a rodent model of pulmonary hypertension. Toxicol Sci 41:209–216.

World Health Organization. 1997. Evaluation of Human Health Risks of PM_{10}. Geneva, Switzerland.

Zeger SL, Dominici F, Samet J. 1999. Harvesting-resistant estimates of air pollution effects on mortality. Epidemiology 10:171–175.

Zelikoff JT, Fang KG, Nadziejko LC, Chen LC, Gordon T, Schlesinger RB, Cohen MD. 1997. Immunomodulation as a mechanism for PM_{10}-induced effects upon host mortality (abstract). In: The Thirteenth Health Effects Institute Annual Conference Program. Health Effects Institute, Cambridge MA.

APPENDIX A. Technical Information

The STL algorithm as applied in this analysis takes each time series (daily deaths, daily $PM_{2.5}$, daily mean temperature, and daily mean dew point) and decomposes it into 3 parts. The first part represents seasonal and other longer-term fluctuations and was fit by applying a LOESS filter to the data with a window of 120 days. The residuals of this process are then filtered again, using a LOESS filter with a second window size (initially set at 15 days). The residuals of this second filter represent fluctuations of less than 15 days. Subtracting these residuals from the residuals of the first filter yields the fluctuations in the original time series that have both long-term and very short-term fluctuations removed. For example, we would decompose daily fluctuations in $PM_{2.5}$ as follows:

$$PM_{2.5} = PM_{2.5long} + PM_{2.5mid} + PM_{2.5short}\,.$$

The subsequent regression analyses are done on the midrange components of each series. LOESS, a nonparametric running line smoother, divides the data for each window into variations that are commensurate with a little more than half the window size or larger and shorter-term variations. This calculation is done by fitting a running regression within each window to estimate the value of the longer frequency component. The regression is weighted with tricubic weights, defined as follows. Let t be the time in days since the beginning of the study, t_{mid} the midpoint of the window (that is, the day for which a smoothed estimate is being computed), and d the half-width of the smoothing window (eg, 60 days for the 120-day smooth that controls for season). Then define u as the fraction of the distance between the midpoint and the end of the window for any observation in the window. That is,

$$u = \frac{(t - t_{mid})}{d}.$$

So u ranges from 0 (at the center of the window) to ±1 at the ends. Then the weights are: $w = (1 - |u|^3)^3$. These weights fall rapidly to zero at the ends of the window and are near 1 for the central 40% of the window (Figure A.1). Once all 4 series were filtered, the following regression was fit:

$$\log(death_{mid} + mean\ [death]) = s(temp_{mid}) + s(dew_{mid}) + PM_{2.5mid}\,.$$

Here, s stands for a nonparametric smooth function, which was used to ensure that nonlinearities in the dependence on weather were adequately modeled. LOESS smoothing was used for this as well, using a span of 50% of the data, which corresponded to approximately 5 *df* for each weather variable. A log-linear model was fit to maintain comparability with the original study. Similarly, temperature and dew point on the concurrent day were used, as in the original paper (Verhoeff et al 1996), and the smoothing window for each weather factor was the same as in the original paper. These conventions maintain maximum comparability to the original results, allowing us to interpret differences in effect size as being due to the exclusion of the very short-term fluctuations in the data from the regressions. To examine the longer windows, the entire process was repeated, using a midscale window of first 30, then 45, and finally 60 days.

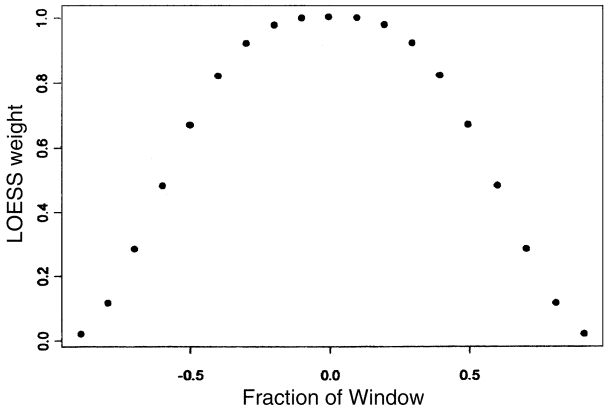

Figure A.1. Weights assigned to points in the LOESS smooth as a function of their distance from the center of the smoothing window. The distance is expressed as a fraction of half the window size and ranges from –1 to 1.

Section 5: Combining Evidence on Air Pollution and Daily Mortality from Twenty Largest US Cities

Francesca Dominici, Jonathan M Samet, and Scott L Zeger

ABSTRACT

Reports over the last decade of association between levels of particles in outdoor air and daily mortality counts have raised concern that air pollution shortens life, even at concentrations within current regulatory limits. Criticisms of these reports have focused on the statistical techniques used to estimate the pollution-mortality relationship and the inconsistency in findings among cities. We have developed analytic methods that address these concerns and combine evidence from multiple locations in order to gain a unified analysis of the data.

Section 5 introduces hierarchical regression models for combining estimates of the pollution-mortality relationship across cities and presents log-linear regression analyses of daily time-series data from the 20 largest US cities as an example. We illustrate this method focusing on health effects of particulate matter less than 10 μm in aerodynamic diameter (PM_{10})* and considering univariate and bivariate analyses with PM_{10} and O_3. In the first stage of the hierarchical model, we estimate the relative mortality rate associated with PM_{10} and O_3 for each of the 20 cities using semiparametric log-linear models. The second stage of the model describes between-city variation in the true relative rates as a function of selected city-specific covariates. We also fit 2 variations of a spatial model with the goal of exploring the spatial correlation of the pollutant-specific coefficients among cities. Finally, to explore the results of considering the 2 pollutants jointly, we fit and compared univariate and bivariate models. All posterior distributions from stage 2 are estimated using Markov chain Monte Carlo (MCMC) techniques. Results appear to be largely insensitive to the specific choice of vague but proper prior distribution. The models and estimation methods are general and can be used for any number of locations and pollutant measurements and have potential application to other environmental agents.

INTRODUCTION

In spite of improvements in measured air quality indicators in many developed countries, the health effects of particulate air pollution remain a regulatory and public health concern. This continued interest is motivated largely by recent epidemiologic studies that have examined both acute and longer-term effects of exposure to particulate air pollution in different cities in the United States and elsewhere in the world (Schwartz 1993; Dockery and Pope 1994; Bascom et al 1996a,b; Korrick et al 1998). Controversial associations have been found using Poisson time-series regression models fit to the data using the generalized estimating equations (Liang and Zeger 1986) or generalized additive models (Hastie and Tibshirani 1990) methods. Model approaches have been questioned (Smith et al 1997; Clyde 1998), although analyses of data from Philadelphia (Kelsall et al 1997; Samet et al 1997) show that the particle-mortality association is reasonably robust to the particular choice of analytic methods from among reasonable alternatives. Past studies have not used a set of communities; most have used data from single locations selected largely on the basis of the data availability or pollution levels. Thus, the extent to which findings from single cities can be generalized is uncertain, and consequently in the National Morbidity, Mortality, and Air Pollution Study (NMMAPS) we select locations from a national sampling frame.

Statistical power of analyses within a single city may be limited by the amount of data for any location. Consequently, pooled analyses can be more informative than analyses of data from a single site about whether or not an association exists, controlling for possible confounders. In addition, a pooled analysis can produce estimates of the parameters at a specific site, which borrow strength from

* A list of abbreviations and other terms appears on page 13.

The National Morbidity, Mortality and Air Pollution Study: Methods and Methodologic Issues, Part I of Health Effects Institute Research Report 94, includes an Investigators' Report, a Preface, a Commentary by the Health Review Committee, and an HEI Statement about the research project. Correspondence concerning this section may be addressed to Dr Francesca Dominici, Department of Biostatistics, Johns Hopkins School of Public Health, 615 North Wolfe Street, Baltimore MD 21205-2179.

Although this document was produced with partial funding by the United States Environmental Protection Agency under Assistance Award R824835 to the Health Effects Institute, it has not been subjected to the Agency's peer and administrative review and therefore may not necessarily reflect the views of the Agency, and no official endorsement by it should be inferred. The contents of this document also have not been reviewed by private party institutions, including those that support the Health Effects Institute; therefore, it may not reflect the views or policies of these parties, and no endorsement by them should be inferred.

data from other locations (DuMouchel and Harris 1983; DuMouchel 1990; Breslow and Clayton 1993).

The main goal of section 5 was to develop a statistical approach that combines information about air pollution–mortality relationships across multiple cities. We illustrate this method with a 2-stage analysis of data from the 20 largest US cities:

1. Given a time series of daily mortality counts in each of 3 age groups, we use generalized additive models (Hastie and Tibshirani 1990) to estimate the relative change in the rate of mortality associated with changes in the air pollution variables, controlling for age-specific longer-term trends, weather, and other potential confounding factors, separately for each city;
2. We then combine the pollution-mortality relative rates across the 20 cities using a Bayesian hierarchical model (Lindley and Smith 1972; Morris and Normand 1992) to obtain an overall estimate and to explore whether some of the geographic variation can be explained by site-specific explanatory variables.

We focus mainly on the second-stage analysis and consider 2 hierarchical regression models—with and without modeling possible spatial correlations—that we refer to as the *baseline model* and the *spatial model*. See Samet and colleagues (1995, 1997) and Kelsall and colleagues (1997) for details on methods for the first-stage analyses, summarized below.

In both models, we assume that the vector of the estimated regression coefficients obtained from the first-stage analysis, conditional on the vector of the true relative rates, has a multivariate normal distribution with mean equal to the true coefficient and a covariance matrix equal to the sample covariance matrix of the estimates. At the second stage of the baseline model, we assume that the city-specific coefficients are independent. In contrast, at the second stage of the spatial model, we allow for a correlation between all pairs of pollutant and city-specific coefficients; these correlations are assumed to decay toward 0 as the distance between the cities increase. Two distance measures are explored.

Next, we briefly describe the database of air pollution, mortality, and meteorologic data from 1987 to 1994 for the 20 US cities in this analysis. We then fit the log-linear generalized additive models to produce relative-rate estimates for each location. The semiparametric regression was conducted 3 times for each pollutant: using the concurrent day (lag 0) pollution values, using the previous day's (lag 1) pollution levels, and using pollution levels from 2 days before (lag 2).

In Pooling Results Across Cities, we present the baseline and the spatial hierarchical regression models for combining the estimated regression coefficients and MCMC methods for model fitting. In particular, we use a Gibbs sampler (Geman and Geman 1984; Gelfand and Smith 1990) for estimating parameters of the baseline model and a Gibbs sampler with a Metropolis step (Hastings 1970; Tierney 1994) for estimating parameters of the spatial model. We next summarize the results, make comparisons between the posterior inferences under the 2 models, and assess the sensitivity of the results to the choice of lag structure and prior distributions.

DATABASES

The analysis database, used for illustrative purposes, includes mortality, weather, and air pollution data for the 20 largest metropolitan areas in the United States for the 7-year period 1987 to 1994 (Figure 1 and Table 1).

The cause-specific mortality data, aggregated at the level of county, were obtained from the National Center for Health Statistics (NCHS). We focused on daily death counts for each site, excluding nonresidents who died in the study site and accidental deaths. Hourly temperature and dew point data for each site were obtained from the EarthInfo CD^2 database. After extensive preliminary analyses that considered various daily summaries of temperature and dew point as predictors, such as daily average, maximum, and 8-hour maximum, we used the 24-hour mean for each day. If there was more than 1 weather station in a city, we took the average of the measurements from all available stations. The PM_{10} and O_3 data were also averaged over all monitors in a county. To protect against outliers, a 10%

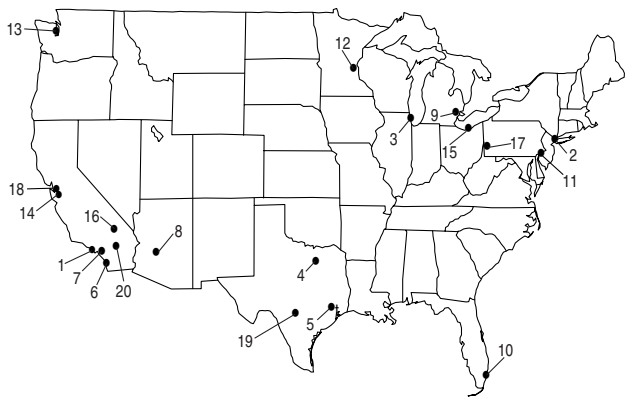

Figure 1. Map showing the 20 cities with the largest populations (including the surrounding county). The cities are numbered from 1 to 20 following the order of Table 1.

Table 1. Summary by Location, County Population, Days with Missing Values, People in Poverty, People 65 Years or Older, Pollutant Levels for O_3 and PM_{10}, and Daily Deaths[a]

Location	Label	Population	P_{missO3} (%)	P_{missPM10} (%)	P_{poverty} (%)	$P_{\geq 65}$ (%)	$\overline{X}_{\text{O3}}$ (ppb)	$\overline{X}_{\text{PM}}$ ($\mu g/m^3$)	$\overline{Y}$
Los Angeles	la	8,863,164	0	80.2	14.8	9.7	22.84	45.98	148
New York	ny	7,510,646	0	83.3	17.6	13.2	19.64	28.84	191
Chicago	chic	5,105,067	0	8.2	14.0	12.5	18.61	35.55	114
Dallas/Fort Worth	dlft	3,312,553	0	78.6	11.7	8.0	25.25	23.84	49
Houston	hous	2,818,199	0	72.9	15.5	7.0	20.47	29.96	40
San Diego	sand	2,498,016	0	82.2	10.9	10.9	31.64	33.63	42
Santa Ana/Anaheim	staa	2,410,556	0	83.6	8.3	9.1	22.97	37.37	32
Phoenix	phoe	2,122,101	0.1	85.1	12.1	12.5	22.86	39.75	38
Detroit	det	2,111,687	36.3	53.9	19.8	12.5	22.62	40.90	47
Miami	miam	1,937,094	1.4	83.4	17.6	14.0	25.93	25.65	44
Philadelphia	phil	1,585,577	0.7	83.1	19.8	15.2	20.49	35.41	42
Minneapolis	minn	1,518,195	100	5.4	9.7	11.6	NA	26.86	26
Seattle	seat	1,507,319	37.3	24.5	7.8	11.1	19.37	25.25	26
San Jose	sanj	1,497,577	0	67.7	7.3	8.6	17.87	30.35	20
Cleveland	clev	1,412,141	41.4	55.6	13.5	15.6	27.45	45.15	36
San Bernardino	sanb	1,412,140	0	81.6	12.3	8.7	35.88	36.96	20
Pittsburgh	pitt	1,336,449	1.3	0.8	11.3	17.4	20.73	31.61	38
Oakland	oakl	1,279,182	0	82.6	10.3	10.6	17.24	26.31	22
San Antonio	sana	1,185,394	0.1	77.1	19.4	9.8	22.16	23.83	20
Riverside	river	170,413	0	81.3	14.8	11.3	33.41	51.99	20

[a] NA = not available; P_{missO3}, P_{missPM10} = percentage of days with missing values; P_{poverty} = percentage of people living in poverty; $P_{\geq 65}$ = percentage of people 65 years or older; $\overline{X}_{\text{O3}}$, $\overline{X}_{\text{PM}}$ = average of pollutant levels for O_3 and PM_{10}; and $\overline{Y}$ = average daily deaths.

trimmed mean was used to average across monitors after correction for yearly averages for each monitor. This yearly correction was appropriate, since long-term trends in mortality were also adjusted in the log-linear regressions (see Kelsall et al 1997 for further details).

CITY-SPECIFIC ANALYSES

Here we summarize the model used to estimate the air pollution–mortality relative rate separately for each location, accounting for age-specific longer-term trends, weather, and day of the week. The core analysis for each city is a log-linear generalized additive model that accounts for smooth fluctuations in mortality that potentially confound estimates of the pollution effect and/or introduce autocorrelation in mortality series. We model daily expected deaths as a function of the pollution levels on the same or immediately preceding days. We build models that include smooth functions of time as predictors as well as the pollution measures to avoid confounding by influenza epidemics, which are seasonal, and by other longer-term factors.

To specify our approach more completely, let y_{at}^c be the observed mortality for each age group $a =$ (<65, 65–75, >75 years) on day t at location c, and x_{at}^c be a $p \times 1$ vector of air pollution variables. Let $\mu_{at}^c = \mathrm{E}(y_{at}^c)$ be the expected number of deaths and $v_{at}^c = \mathrm{var}(y_{at}^c)$. We use a log-linear model $\log \mu_{at}^c = x_{at}^{c\prime}\beta^c$ for each city c, allowing the mortality counts to have variances v_{at}^c that may exceed their means (ie, be overdispersed) with the overdispersion parameter ϕ^c also varying by location so that $v_{at}^c = \phi^c \mu_{at}^c$.

To protect the pollution relative rates β^c from confounding by longer-term trends due, for example, to changes in health status, changes in the sizes and characteristics of populations, seasonality, and influenza epidemics, and to account for any additional temporal

correlation in the count time-series, we estimated the pollution effect using only shorter-term variations in mortality and air pollution. To do so, we partial out the smooth fluctuations in the mortality over time by including arbitrary smooth functions of calendar time $S^c(time, \lambda)$ for each city. Here, λ is a smoothness parameter that we prespecified, based on prior epidemiologic knowledge of the time scale of the major possible confounders, to have 7 degrees of freedom (*df*) per year of data so that little information from time scales longer than approximately 2 months is included when estimating β^c. This choice largely eliminates expected confounding from seasonal influenza epidemics and from longer-term trends that result from changing medical practice and health behaviors, while it retains as much unconfounded information as possible. We also controlled for age-specific longer-term temporal variations in mortality, adding a separate smooth function of time with 8 *df* for each age group.

To control for weather, we also fit smooth functions of the same-day temperature ($temp_0$), average temperature for the 3 previous days ($temp_{1-3}$), each with 6 *df*, and the analogous functions for dew point (dew_0, dew_{1-3}), each also with 3 *df*. Since there were missing values of some predictor variables on some days, we restricted analyses to days with no missing values across the full set of predictors.

In summary, we fit the following log-linear generalized additive model (Hastie and Tibshirani 1990) to obtain the estimated pollution relative rate β^c and the sample covariance matrix V^c_β at each location:

$$\begin{aligned}\log \mu^c_{at} = {} & x^{c'}_{at}\beta^c + \gamma^c DOW + S^c_1(time, 7/\text{year}) \\ & + S^c_2(temp_0, 6) + S^c_3(temp_{1-3}, 6) \\ & + S^c_4(dew_0, 3) + S^c_5(dew_{1-3}, 3) \\ & + \text{intercept for age group } a \\ & + \text{time (8 } df\text{) for age group } a, \end{aligned} \quad (1)$$

where *DOW* are indicator variables for day of week. Samet and colleagues (1995, 1997) and Kelsall and colleagues (1997) give additional details about choices of functions used to control for longer-term trends and weather. Alternative modeling approaches that consider different lag structures of the pollutants and of the meteorological variables have been proposed (Davis et al 1996; Smith et al 1997, 1998). More general approaches that consider nonlinear modeling of the pollutant variables have been discussed by Smith and colleagues (1997).

Because the functions $S^c(x, \lambda)$ are smoothing splines with fixed λ, the semiparametric model described above has a finite-dimensional representation. Hence, the analytic challenge is to make inferences about the joint distribution

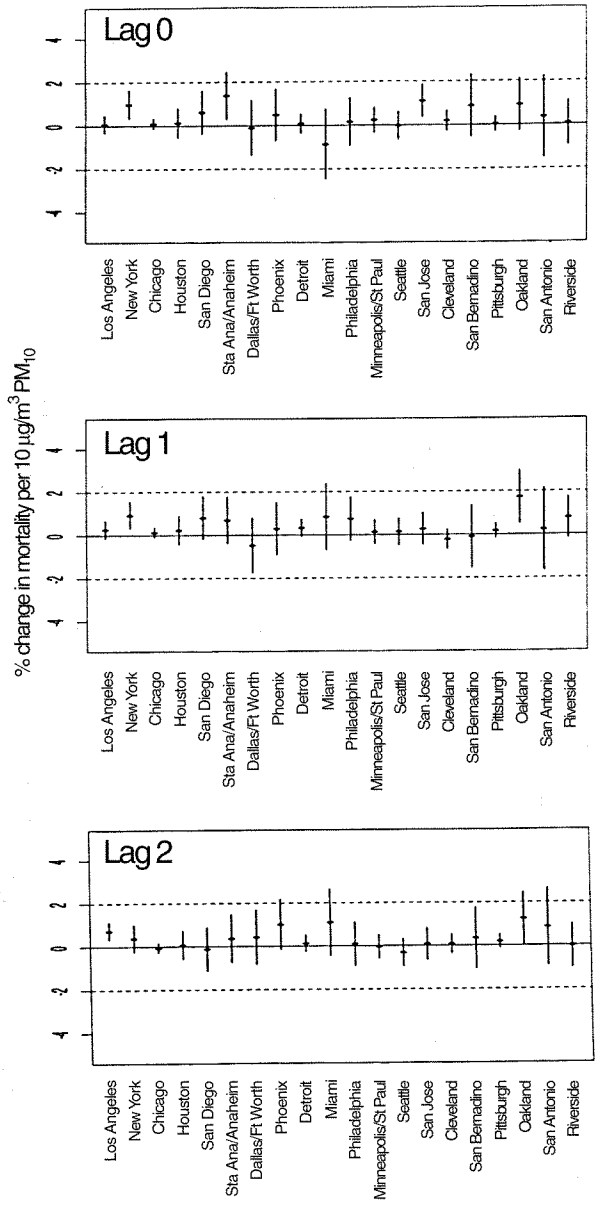

Figure 2. Univariate model for PM_{10}. Results of regression models for the 20 cities by selected lag: $\hat{\beta}^c$ and 95% CIs of $\hat{\beta}^c \times 1{,}000$ for PM_{10}. Cities are presented in decreasing order by population living within their county limits. The vertical scale can be interpreted as the percentage increase in mortality per 10 $\mu g/m^3$ increase in PM_{10}. The results are reported using the concurrent day (lag 0) pollution values to predict mortality, using the previous day's (lag 1) pollution levels, and using pollution levels from 2 days before (lag 2).

of the β^c in the presence of finite-dimensional nuisance parameters, which we refer to as η^c.

We separately estimated 3 semiparametric regressions for each pollutant with the concurrent day (lag 0), prior day (lag 1), and 2 days prior (lag 2) pollution predicting mortality. The estimates of the coefficients and their 95% confidence intervals (CIs) for PM_{10} entered independently in the model and for PM_{10} adjusted by O_3 are shown in Figures 2 and 3. Cities are presented in decreasing order by the size of their populations. The figures show substantial between-location variability in the estimated relative rates, suggesting that combining evidence across cities would be a natural approach to explore possible sources of heterogeneity and to obtain an overall summary of the degree of association between pollution and mortality. To add flexibility in modeling the lagged relationship of air pollution with mortality, we could have used distributed lag models instead of treating the lags separately. Although desirable, this was not easily implemented because many cities have PM_{10} data available only every sixth day.

To test whether the log-linear generalized additive model (1) has taken appropriate account of the time dependence of the outcome, we have calculated, for each city, the autocorrelation function of the standardized residuals. Figure 4 displays the 20 autocorrelation functions; they are centered near 0 and range between –0.05 and 0.05, confirming that the filtering has removed the serial dependence.

We also examined the sensitivity of the pollution relative rates to the *df* used in the smooth functions of time, weather, and seasonality by halving and doubling each of them. The relative rates changed very little as these parameters were varied over this fourfold range (data not shown).

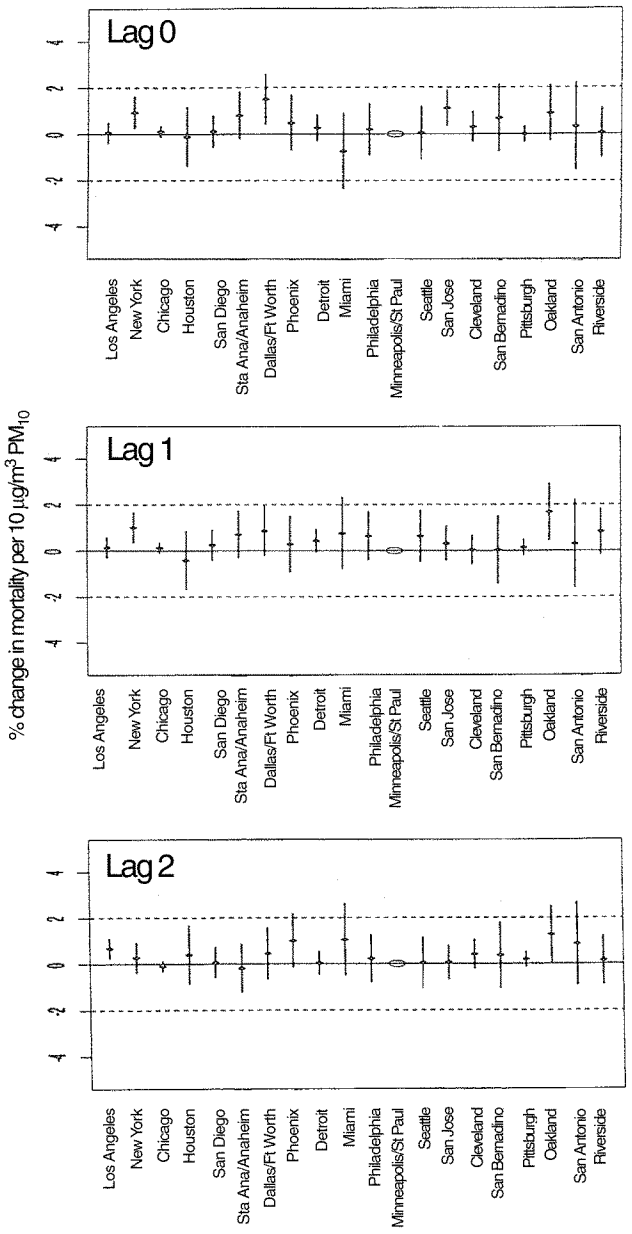

Figure 3. Bivariate model for PM_{10} adjusted by O_3. Results of regression models for the 20 cities by selected lag: $\hat{\beta}^c$ and 95% CIs of $\hat{\beta}^c \times 1{,}000$ for PM_{10} adjusted by O_3. Cities are presented in decreasing order by population living within their county limits. The empty circle placed at Minneapolis represents the absence of the O_3 data in this city. The vertical scale can be interpreted as the percentage increase in mortality per 10 µg/m^3 increase in PM_{10}. The results are reported using the concurrent day (lag 0) pollution values to predict mortality, using the previous day's (lag 1) pollution levels, and using pollution levels from 2 days before (lag 2).

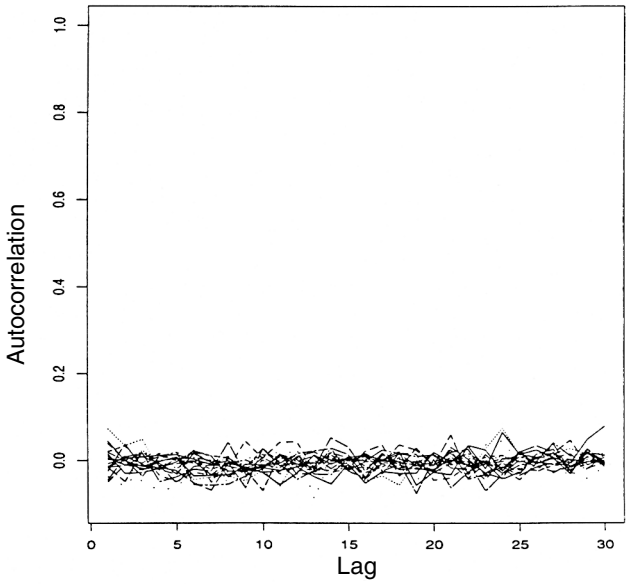

Figure 4. Plots of city-specific autocorrelation functions of standardized residuals r_t, where $r_t(Y_t - \hat{Y}_t) / \sqrt{(\hat{Y}_t)}$ and $\hat{Y}_t$ are the fitted values from log-linear generalized additive model (1).

POOLING RESULTS ACROSS CITIES

Here we present hierarchical regression models designed to pool the city-specific pollution relative rates across cities to obtain a summary value. Hierarchical regression models provide a flexible approach to the analysis of multilevel data. In this context, the hierarchical approach provides a unified framework for making estimates of the city-specific pollution effects, the overall pollution effect, and the within-cities and between-cities variation of the city-specific pollution effects.

Results of several applied analyses using hierarchical models have been published. Examples include models for the analysis of longitudinal data (Gilks et al 1993), spatial data (Breslow and Clayton 1993), and health care utilization data (Normand et al 1997). Other modeling strategies for combining information in a Bayesian perspective are provided by DuMouchel (1990), Skene and Wakefield (1990), Smith and colleagues (1995), and Silliman (1997). Recently, spatiotemporal statistical models with applications to environmental epidemiology have been proposed by Wickley and colleagues (1997), Meiring and colleagues (1997), and Wakefield and Morris (1998).

After the following overview of our modeling strategy, we consider 2 hierarchical regression models with and without modeling of the possible spatial autocorrelation among the β, which we refer to as the *baseline* and *spatial* models, respectively.

MODELING APPROACH

The modeling approach comprises 2 stages. At the first stage, we use the log-linear generalized additive model (1):

$$y_t^c \mid \beta^c, \eta^c \sim \text{Poisson}\left[\mu_t\left(\beta^c, \eta^c\right)\right]$$

where $y_t^c = (y_{<65t}, y_{65-75t}, y_{>75t})$. The parameters of scientific interest are the mortality relative rates, β^c, which for the moment are assumed not to vary across the 3 age groups within a city. The vector η^c of the coefficients for all the adjustment variables, including the splines in the semiparametric log-linear model, is a finite-dimensional nuisance parameter.

The second stage of the model describes variation among the β^c across cities. We regress the true relative rates on city-specific covariates, z^c, to obtain an overall estimate and to explore the extent to which the site-specific explanatory variables explain geographic variation in the relative risks. In epidemiologic terms, the covariates in stage 2 are possible effect modifiers. More specifically, we assume:

$$\beta^c \mid \alpha, \Sigma \sim N_p\left(z^c\alpha, \Sigma\right)$$

where p is the number of pollutant variables that enter simultaneously in model (1). Here the parameters of scientific interest are the vector of the regression coefficients, α, and the overall covariance matrix, Σ. Because there is little or no interest in the η^c in the current scientific context, we assume that these nuisance parameters are independent; that is, we do not use data from the other locations for the purpose of improving estimation of the remaining location.

Our goal is to make inferences about parameters of interest—the β^c, α, and Σ—in the presence of nuisance parameters η^c. To estimate an exact Bayesian solution to this pooling problem, we would analyze the joint posterior distributions of the parameters of interest as well as of the nuisance parameters, and then integrate over the η^c dimension to obtain the marginal posterior distributions of the β^c. Although possible, the computations become extremely laborious and are not practical for either this analysis or a planned model with 100 or more cities.

Given the large sample size at each city (T ranges from 550 to 2,550 days), accurate approximations to the posterior distribution can be obtained using the normal approximation of the likelihood (Le Cam and Yang 1990). Therefore, if the joint likelihood $p(y^c \mid \beta^c, \eta^c)$ has a multivariate normal distribution with mean equal to the maximum likelihood estimates of β^c and η^c and covariance matrices V_β^c and V_η^c, then by definition $p(y^c \mid \beta^c)$ has multivariate normal distribution with mean $\hat{\beta}^c$ and covariance matrix $V_{\hat{\beta}}^c$. We then replace the first stage of the model with a normal distribution, with mean and variance equal to the maximum likelihood estimates of the parameter. Recently it has been shown that the strategy based on the normal approximation of the likelihood gives an alternative 2-stage model that well approximates the original model and leads to more efficient simulation from the posterior (Daniels and Kass 1998).

To check whether the inferences based on the normal approximation of the likelihood are proper, we have compared our approach with the implementation of the full MCMC approach for a few cities with sample sizes ranging from 2,000 in Pittsburgh to 545 in Riverside. Figure 5 shows the histogram of samples for Riverside from $p(\beta^c \mid data)$—obtained implementing a Gibbs sampler that simulates from $p(\beta^c \mid \eta^c, data)$ and $p(\eta^c \mid \beta^c, data)$ and approximate $p(\beta^c \mid data)$ by $\int p(\beta^c, \eta^c \mid data) d\eta^c$—with samples from $N(\hat{\beta}^c, V_{\hat{\beta}}^c)$ (the solid line in the figure). The 2 distributions are very similar.

BASELINE MODEL

Let $\beta^c = [\beta^c_{PM10}, \beta^c_{O3}]'$ be the relative rate associated with PM_{10} and O_3 at city c. We consider the following hierarchical model:

$$\hat{\beta}^c \mid \beta^c \sim N_2(\beta^c, V^c_\beta)$$

$$\beta^c_{PM10} = z^{c'}_{PM10}\alpha_{PM10} + \varepsilon^c_{PM10} \qquad (2)$$

$$\beta^c_{O3} = z^{c'}_{O3}\alpha_{O3} + \varepsilon^c_{O3}$$

$$\varepsilon^c \mid \Sigma \sim N_2(0, \Sigma)$$

where $z^c_{PM10} = [1, P^c_{poverty}, P^c_{>65}, \overline{X}^c_{PM10}]'$, $z^c_{O3} = [1, P^c_{poverty}, P^c_{>65}, \overline{X}^c_{O3}]'$, α_{PM10} and α_{O3} are 4×1 vectors, and finally $\varepsilon^c = [\varepsilon^c_{PM10}, \varepsilon^c_{O3}]'$ c = 1, ..., 20. This model specification allows dependence between the relative rates associated with PM_{10} and O_3 but implies independence between the relative rates of city c and c'.

Under this model, the true PM_{10} and O_3 relative rates in city c are regressed on predictor variables, including, for example, the percentage of people in poverty ($P^c_{poverty}$) and the percentage of people older than 65 years ($P^c_{>65}$), and on the average of the daily values of PM_{10} and O_3 over the period 1987 to 1994 in location c, $\overline{X}^c_{PM10}$, $\overline{X}^c_{O3}$. If we center the predictors about their means, the intercepts $\alpha_{0,PM10}$ and $\alpha_{0,O3}$ can be interpreted as overall effects for a city with mean predictors. A simple pooled estimate of the pollution effect is obtained by setting all covariates to 0. To compare the consequences of considering 2 pollutants independently and jointly in the model, we fit a baseline-univariate model—that is, Σ assumed diagonal—and a baseline-bivariate model—that is, Σ assumed to have nonzero off-diagonal elements.

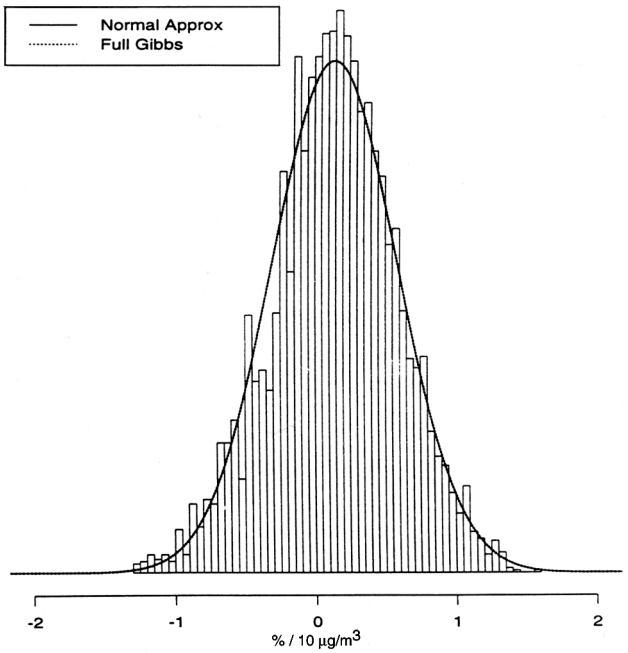

Figure 5. Comparison between the normal approximation of the likelihood of β^c and the marginal posterior distribution of β^c. The solid line represents the normal density $N(\hat{\beta}^c, V^c_\beta)$ where $\hat{\beta}^c$ and V^c_β are the maximum likelihood estimates of a semiparametric Poisson regression model. The histogram represents the marginal posterior distribution of β^c obtained by implementing a full Gibbs sampler for the parameter of interest β^c and for the coefficients of the natural cubic splines η^c.

Inference on the parameters $\alpha = [\alpha_{PM10}, \alpha_{0,O3}]'$ and Σ represents a synthesis of the information from the 20 cities; for example, the parameters α_{0j}, $[\Sigma]_{jj}$ $j = PM_{10}, O_3$, determine the overall level and the variability of the relative change in the rate of mortality associated with changes in the jth pollutant level on average over all the cities.

The Bayesian formulation is completed by specifying dispersed but proper baseline prior distributions, and then supplementing the baseline analysis with additional sensitivity analysis. A priori, we assume that the joint prior is the product of the marginals for α and Σ. The following baseline prior specifications for the marginals were used:

Overall relative rates $\quad \alpha \sim N_{p(k+1)}(m, V_\alpha)$

Overall covariance matrix $\quad \Sigma \sim IW_p(df, D)$

where p and k denote the number of pollutant variables entering simultaneously in the model and k the number of city-specific covariates, respectively. We select m equal to a vector of 0s, V_α equal to a diagonal matrix, with diagonal elements equal to 100, $df = 3$, and D a diagonal matrix with diagonal elements equal to 3. These prior hyperparameters lead prior 95% support to the overall effect, the city-specific effects, and the correlation between the PM_{10} and the O_3 relative rates equal to (–10, 10), (–4, 4), and (–0.85, 0.85), respectively. This prior specification has been selected because it does not impose too much shrinkage of the study-specific parameters toward their overall means, while at the same time specifying a reasonable range for the unknown parameters a priori.

Given these prior assumptions, we can draw inferences on the unknown parameters using the posterior distribution

$$p(\beta^1, ..., \beta^{20}, \alpha, \Sigma \mid \hat{\beta}^1, ..., \hat{\beta}^{20}, V^1_\beta, ..., V^{20}_\beta). \qquad (3)$$

To do this, we implement an MCMC algorithm with a block Gibbs sampler (Gelfand and Smith 1990) in which the unknowns are partitioned into the following groups: β^c, α, and Σ. Each group is sampled in turn, given all others. The full conditional distributions are all available in closed form. Their derivation is routine (Bernardo and Smith

1994) and not detailed here. Because of the normality assumptions at the first and second stages of the hierarchical model, computation of the posterior distributions of all the unknowns under an univariate model can be performed via direct simulation following the factorization:

$$p(\beta^1, \ldots, \beta^{20}, \alpha, \sigma^2 | data) = p(\sigma^2 | data) p(\alpha | \sigma^2, data) \prod_{c=1}^{20} p(\beta^c | \alpha, \sigma^2, data)$$

The first step, simulating σ^2, can be performed numerically (using the inverse cumulative distribution function method, for example). The second and third steps can be done easily by sampling from normal distributions. This strategy can be conveniently implemented only for the univariate baseline model.

SPATIAL MODEL

The assumption of independence of the city-specific coefficients made in the baseline model can be relaxed to a more general model in which the correlation between β^c and $\beta^{c'}$ decays as either a smooth or step function to 0 as the distance between the 2 cities, c and c', increases. Here we consider a hierarchical model in which the inferences allow for the possible spatial correlation among the β^c. We consider only univariate models because of the small number of cities; extension to multivariate models is straightforward but requires a larger data set.

At the second stage of the spatial model, we assume that there is systematic variation in the air pollution–mortality relationship from pollutant to pollutant as specified in the baseline model (2). We express the degree of similarity of the relative rates in locations c and c', as a function of an (arbitrary) distance between c and c', by assuming $\rho(c, c') = \mathrm{Corr}(\beta^c, \beta^{c'}) = \exp(-\theta d(c, c'))$. We consider 2 distance measures, the Euclidean distance between the cities c and c' in the longitude and latitude coordinates, and a step function such that $d(c,c) = 1$ if location c and c' are within a common region and $d(c, c') = \infty$ if not. Under this assumption, the correlation of 2 cities within the same region is $e^{-\theta}$, and the correlation for cities from different regions is 0. Alternative definitions of distance can be incorporated easily into the model as appropriate.

The spatial model with $(1, \infty)$ distance can be also specified as a 3-stage hierarchical model where the first stage is as the baseline model (2), the second stage describes the heterogeneity of the estimates within regions, and the third stage describes the heterogeneity of the estimates across regions. For this regional model, we have clustered the 20 cities in the following 3 regions: East, South, and West US.

The spatial model specification is completed with the elicitation of the prior distribution. For α and Σ we choose the same prior specified for the baseline model. For the parameter θ under the spatial model with Euclidean distance, we choose a log-normal prior with mean 1 and standard deviation 1.2. This specification leads to a prior distribution of the correlation $\rho(c, c')$ of the closest cities having mean 0.56 (95% interval, 0 to 0.96), and prior distributions of all the other pairs c, c' more concentrated at 0 as the distance between cities increases.

In the spatial model, the full conditionals for the β^c, α, and Σ are all available in closed form. In contrast, to sample from the full conditional distribution of θ, we used a Metropolis-Hastings algorithm with a Gamma proposal distribution having mean equal to the current value of θ and fixed variance. The spatial model with step distance can be more efficiently sampled with a block Gibbs sampler because the full conditional distributions of all the unknown parameters are available in closed form.

RESULTS

We ran the Gibbs sampler for 3,000 iterations for both the baseline and the spatial models, ignoring the first 100. The autocorrelation, computed from a random sample of the $\alpha_{0,\,PM10}$, was negligible at lag 5 so we sampled every fifth observation for posterior estimation. The acceptance probabilities for the Metropolis-Hastings algorithm averaged between 0.3 and 0.5. Convergence diagnosis has been performed by implementing Raftery and Lewis (1992) methods in Convergence Diagnostics and Output Analysis (CODA) (Best et al 1995), which reports the minimum number of iterations, *Nmin*, needed to estimate the variable of interest with an accuracy of ±0.005 and with probability of attaining this degree of accuracy equal to 0.95. *Nmin* ~ 2,000 were proposed.

Figure 6 summarizes results of the pooled analyses under the univariate-baseline model. This figure displays the posterior distributions of city-specific regression coefficients β^c associated with changes in PM_{10} measurements for the 20 cities at the current-day, 1-day, and 2-day lags. The marginal posterior distribution of the overall effect ($\alpha_{0,\,PM10}$) is displayed at far right. Cities are ordered by the decreasing size of their populations. At the current day, the highest relative rate for the PM_{10} variable occurs at Santa Ana/Anaheim with 1.08% increase in mortality (95% interval, 0.20 to 1.97) per 10 $\mu g/m^3$ increase in PM_{10}. Overall, we find that a

10 μg/m^3 increase of PM_{10} is associated with an estimated 0.34% increase in mortality (95% interval, –0.08 to 0.78).

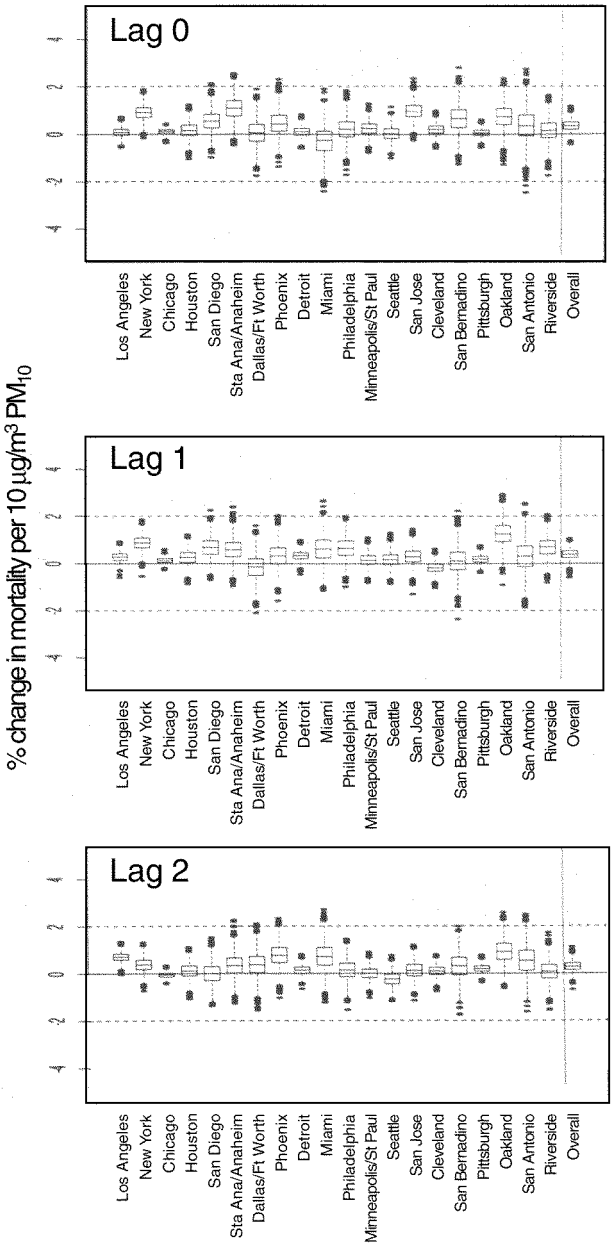

Figure 6. Univariate model for PM_{10}. Results of pooled analyses (PM_{10} entered independently in the model). Boxplots of samples from the posterior distributions of city-specific regression coefficients, β^c, associated with changes in PM_{10} measurements. For comparison, samples from the marginal posterior distribution of the corresponding overall effects are displayed at far right. The vertical scale can be interpreted as the percentage increase in mortality per 10 μg/m^3 increase in PM_{10}. The results are reported using the concurrent day (lag 0) pollution values to predict mortality, using the previous day's (lag 1) pollution levels, and using pollution levels from 2 days before (lag 2).

Figure 7 summarizes results of the pooled analyses under the bivariate-baseline model. When PM_{10} and O_3 are combined in the same model, we estimate that 10-unit increments in PM_{10} adjusted by O_3 are associated with

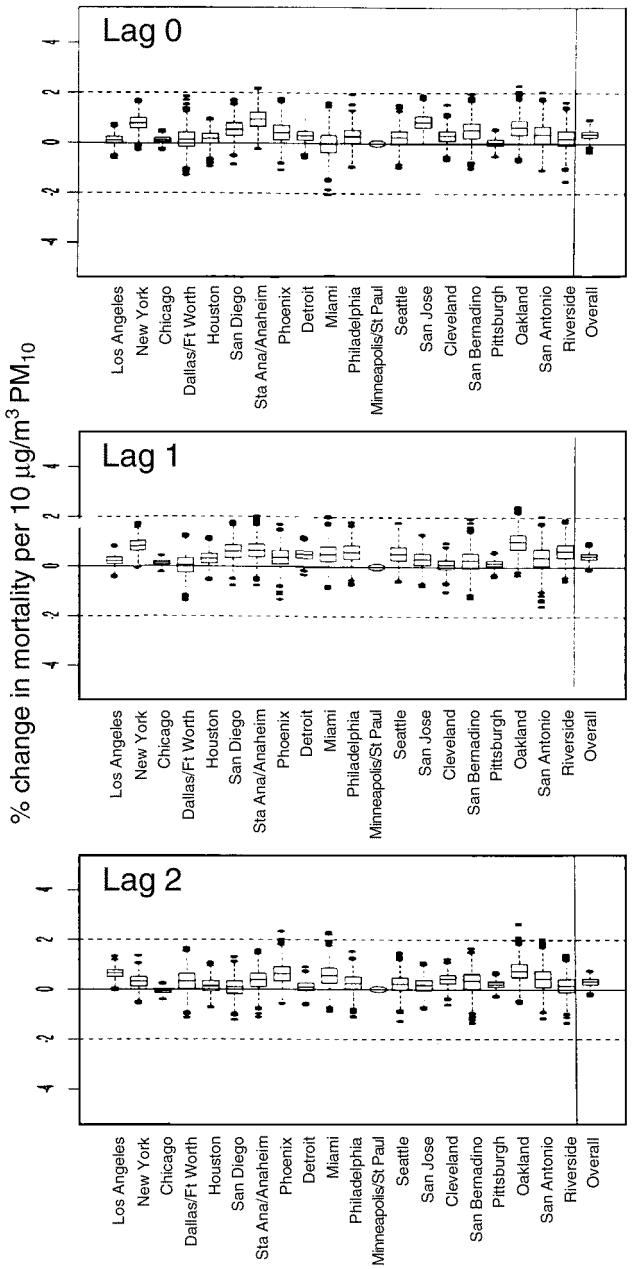

Figure 7. Bivariate model for PM_{10} adjusted for O_3. Results of pooled analyses under the baseline bivariate model (PM_{10} and O_3 entered simultaneously in the model). Boxplots of samples from the posterior distributions of city-specific regression coefficients, β^c, associated with changes in PM_{10} adjusted by O_3 measurements. For comparison, samples from the marginal posterior distribution of the corresponding overall effects are displayed at far right. The vertical scale can be interpreted as the percentage increase in mortality per 10 μg/m^3 increase in PM_{10}. The results are reported using the concurrent day (lag 0) pollution values to predict mortality, using the previous day's (lag 1) pollution levels, and using pollution levels from 2 days before (lag 2).

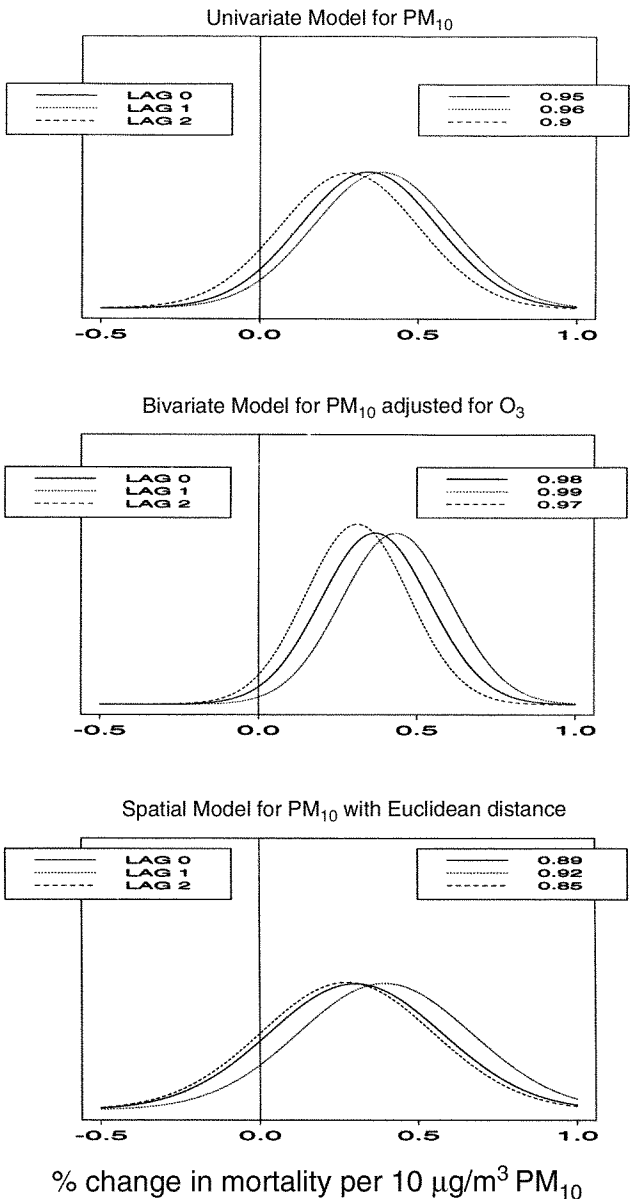

Figure 8. Pooled analyses under the baseline univariate, baseline bivariate, and spatial models. Marginal posterior distributions of the overall effects, $\alpha_{0,\,PM10}$, for different lags. At the top right are specified the posterior probabilities that the overall effects are larger than 0.

mortality increases of 0.37% (95% interval, 0.04 to 0.71) increase in mortality.

The marginal posterior distribution of the overall regression effect combines and synthesizes the information from the 20 locations. Figure 8 shows the marginal posterior distributions of the overall pollution relative rates at the current day, 1-day, and 2-day lags obtained from the baseline-univariate, the baseline-bivariate, and the spatial models. At the top right the posterior probabilities that the overall effects are larger than 0 for each lag-specification are summarized. In univariate and bivariate analyses, we found significant effects of PM_{10}.

Results of the adjusted analyses under the univariate-baseline model are shown in Table 2. Here we summarize the posterior means and the 95% posterior support intervals for the relationship between the mean of the city-specific coefficients and the percentage in poverty, the percentage of people older than 65, and the mean level of the pollutant. None of these variables were found to predict the PM_{10} relative rate.

An interaction of the pollution effects and age can be detected by the coefficient of the variable $P_{>65}$ in the second-stage regression model. A more direct approach is to estimate a separate pollution relative rate for each age stratum in the first-stage log-linear models and then to pool the trivariate vector $(\hat{\beta}_{<65}, \hat{\beta}_{65-75}, \hat{\beta}_{>75})$ across cities. When we did this, the estimate of the overall effect of PM_{10} for persons older than 75 was the largest with posterior mean 0.46 (95% interval, 0.12 to 0.82), but there was not a strong trend in the pollution relative rates with age as was suggested by the second-stage regression results in Table 2.

The variability of the regression coefficients, on average, over all the locations is captured by the matrix Σ. Marginal posterior means and 95% posterior support intervals are summarized in Table 3. A large diagonal element signifies large variability over cities in the corresponding coefficient, while a large off-diagonal element signifies strong correlation between the PM_{10} and O_3 coefficients. Table 3 shows the results. Under the baseline bivariate model, the standard deviation of the true coefficients across cities is estimated to be 0.60 (95% interval, 0.42 to 0.84), which is about twice as large as the overall estimate of the pollution effect. Hence, the variability in PM_{10} coefficient is nonnegligible. The posterior distribution of the off-diagonal elements of Σ indicates a negative mean correlation between the effects of the 2 pollutants, but with a large standard deviation.

From the posterior samples of θ in the spatial model, we can easily calculate the marginal posterior distributions of the correlation coefficient $\rho(c, c') = \exp(-\theta d(c, c'))$ for each distance $d(c, c')$. For the 2 closest cities, the posterior mean and interquartile range (IQR) of the correlation between β^c and $\beta^{c'}$ is 0.73 (0.51 to 0.88) for PM_{10}. Under the regional model, with distance equal to a step function, the posterior mean and interquartile range (IQR) of the within-region correlation of the city-specific relative rates, $e^{-\theta}$, is 0.73 (0.66 to 0.86), suggesting that the adverse health effects of PM_{10} on mortality are more similar for locations belonging to the same region than for locations belonging to 2 different regions. The posterior means and IQR for the

Table 2. Results of Adjusted Second-Stage Analyses Under Baseline Univariate Model[a]

City-Specific Covariates	Lag 0	Lag 1	Lag 2
Overall PM_{10}	0.350 (−0.099, 0.815)	0.380 (−0.078, 0.828)	0.236 (−0.183, 0.667)
$P_{poverty}$ (%)	−0.015 (−0.141, 0.112)	0.011 (−0.113, 0.131)	0.040 (−0.074, 0.160)
$P_{>65}$ (%)	−0.050 (−0.224, 0.123)	0.027 (−0.144, 0.207)	−0.005 (−0.175, 0.169)
$\bar{X}^c_{PM10}$ ($\mu g/m^3$)	−0.003 (−0.065, 0.055)	−0.004 (−0.065, 0.058)	−0.006 (−0.064, 0.053)

[a] PM_{10} is entered independently in the model. Posterior means (95% posterior support intervals) of coefficients for the relationship between true relative rate (β^c), percentage of people in poverty ($P_{poverty}$), percentage of people older than 65 ($P_{>65}$), and mean level of pollutant ($\bar{X}^c_{PM10}$). Results are reported using concurrent day's (lag 0) pollution values to predict mortality, using previous day's (lag 1) pollution levels, and using pollution levels from 2 days previously (lag 2).

Table 3. Posterior Means (and 95% Support Intervals) of Elements of Σ Under Three Models (Baseline Bivariate, Baseline Univariate, and Spatial)[a]

Model	Std of PM_{10} Effects	Std of O_3 Effects	Correlation Between PM_{10} and O_3 Effects
Baseline Bivariate	0.60 (0.42, 0.84)	0.84 (0.51, 1.29)	−0.12 (−0.60, 0.42)
Baseline Univariate	0.87 (0.65, 1.18)	1.04 (0.75, 1.41)	
Spatial	0.90 (0.65, 1.24)	1.05 (0.75, 1.39)	

[a] Std of PM_{10} effects = standard deviation across locations of β^c_{PM10}; Std of O_3 effects = standard deviation across locations of β^c_{O3}; correlation between PM_{10} and O_3 effects = correlation between β^c_{PM10} and β^c_{O3}.

regional effects β^{EAST}, β^{SOUTH}, and β^{WEST} are 0.20 (0 to 0.44), 0.1 (–0.2 to 0.39), 0.52 (0.27 to 0.75), suggesting that the adverse health effects of PM_{10} on mortality in the West US was larger than in the East and South US.

We have assessed the robustness of the results with respect to choices of the model (univariate, bivariate, spatial), of the lag-structure (lag 0, lag 1, lag 2), and of the prior distributions. Our sensitivity analysis compared 27 alternative scenarios (3 for model choice, 3 for lag structures, and 3 for prior distributions). For these scenarios we compared the posterior probability that the overall effect of the PM_{10} is larger than 0. The consequences of these choices are shown in Table 4. Significant effects of PM_{10} on total daily mortality were observed in all 3 models (weaker under a spatial model with 2 days prior—lag 2—pollution predicting mortality). When both pollutants were included in the model, adverse effects of PM_{10} become stronger. Spatial analyses slightly attenuate the effects.

Table 4. Posterior Probabilities That Overall Effects of PM_{10} Are Larger Than 0 by Lag and by 3 Prior Distributions Under 3 Models (Baseline Univariate, Baseline Bivariate, and Spatial)[a]

	Prior 1			Prior 2			Prior 3		
Model	Lag 0	Lag 1	Lag 2	Lag 0	Lag 1	Lag 2	Lag 0	Lag 1	Lag 2
Baseline Univariate	0.95	0.96	0.90	0.93	0.95	0.93	0.95	0.96	0.90
Baseline Bivariate	0.98	0.99	0.97	0.99	0.98	0.97	0.98	0.99	0.97
Spatial	0.89	0.92	0.85	0.91	0.94	0.84	0.88	0.89	0.84

[a] The 3 prior specifications lead to the following 95% support intervals of the overall effects and the city-specific effects, and to the following 75% support intervals for the spatial correlation for the relative rates of the 2 closest cities: Prior 1: (–20, 20), (–4, 4), (0.16, 0.8); Prior 2: (–4, 4), (–4, 4), (0.16, 0.8); and Prior 3: (–4, 4), (–4, 4), (0.64, 0.9).

DISCUSSION

We have developed a statistical model for obtaining a national estimate of the effect of urban air pollution on daily mortality and illustrated its use. Because estimation of a national pollution-relative rate is the primary objective of NMMAPS, we used a 2-stage approach that allowed the modeling effort to focus on combining information across cities. In the first stage, we used a log-linear regression to estimate a pollution-relative rate for each city while controlling for the city-specific longer-term time trends and weather effects. Because we had no specific scientific interest in the time or weather effects, we made no effort to impose modeling assumptions to enable us to borrow strength across cities when estimating the effects on mortality of these variables.

In the second stage, we regressed the true relative rates on city-specific covariates to obtain an overall estimate and to estimate the variation among the coefficients across cities. We then generated posterior estimates of the overall pollution effect and of the city-specific effects using MCMC methods. We used 2 models for combining information across cities. The first treated relative rates from different cities as independent of one another. The second allowed the possibility of geographic correlation among the true coefficients.

Although only a first step, the modeling described here establishes a basis for carrying out national surveillance for effects of air pollution and weather on public health. The analyses can be easily extended to studies of cause-specific mortality and other pollutants. Monitoring efforts using models such as the one described here are appropriate given the important public health questions that they can address and the considerable expense to government agencies for collecting the information that forms the basis for this work.

An alternative modeling strategy would have been to use a single large MCMC method to estimate simultaneously the parameters in the log-linear models within each city, the overall estimate of the pollutant, and all of the nuisance parameters. Such a model would borrow strength across cities to obtain more precise estimates of the nuisance functions for each city. This type of approach would be necessary if there were limited information about the nuisance parameters within each city as, for example, in the Neyman and Scott (1960) problem. As this was not the case in our investigation, we focused the modeling and computing effort on combining city-specific relative rate estimates to obtain a national average relative rate.

If the likelihood function for the pollution relative rate and the nuisance parameters is well approximated by a Gaussian distribution, then our approach will give a close approximation to the posterior distribution from an MCMC that simulated both the parameters of interest and the nuisance parameters. We compared the marginal posterior of the β^c obtained using a full MCMC with our normal approximation for a few cities; they were indistinguishable.

This new hierarchical method will be applied to data from 90 US cities. The 20-city example shows the potential for this approach to provide a summary estimate of the effect of particulate matter while controlling for other pollutants and considering differences in city characteristics.

ACKNOWLEDGMENTS

The authors are grateful to Dr Giovanni Parmigiani and to Dr Frank Curriero for comments on an earlier draft, and to Mr Ivan Coursac and Ms Jing Xu for development of the database.

REFERENCES

Bascom R, Bromberg PA, Costa DA, Devlin R, Dockery DW, Frampton MW, Lambert W, Samet JM, Speizer FE, Utell M. 1996a. Health effects of outdoor air pollution. Part 1. Am J Respir Crit Care Med 153:3–50.

Bascom R, Bromberg PA, Costa DA, Devlin R, Dockery DW, Frampton MW, Lambert W, Samet JM, Speizer FE, Utell M. 1996b. Health effects of outdoor air pollution. Part 2. Am J Respir Crit Care Med 153:477–498.

Bernardo JM, Smith AFM. 1994. Bayesian Theory. John Wiley & Sons, New York NY.

Best NG, Cowles MK, Vines K. 1995. CODA: Convergence Diagnostics and Output Analysis Software for Gibbs Sampling Output, Version 0.30. Technical Report. MRC Biostatistics Unit, University of Cambridge, Cambridge, England.

Breslow NE, Clayton DG. 1993. Approximate inference in generalized linear mixed models. J Am Stat Assoc 88:9–25.

Clyde M. 1998. Bayesian model averaging and model search strategies. In: Bayesian Statistics 6. Proceedings of the Sixth Valencia International Meeting, Valencia, Spain. Oxford University Press, Oxford, England.

Daniels M, Kass R. 1998. A note on first-stage approximation in two-stage hierarchical models. Sankhya, Series B 60:19–30.

Davis J, Sacks J, Saltzman N, Smith R, Styer P. 1996. Airborne Particulate Matter and Daily Mortality in Birmingham, Alabama. Technical Report. National Institute of Statistical Science, Research Triangle Park NC.

Dockery DW, Pope CA III. 1994. Acute respiratory effects of particulate air pollution. Annu Rev Public Health 15:107–132.

DuMouchel WM. 1990. Bayesian meta-analysis. In: Statistics: Textbooks and Monographs, Vol 104 (Berry DA, ed). Marcel Dekker, New York NY.

DuMouchel WH, Harris JE. 1983. Bayes methods for combining the results of cancer studies in humans and other species. J Am Stat Assoc 78:293–308.

Gelfand AE, Smith AFM. 1990. Sampling based approaches to calculating marginal densities. J Am Stat Assoc 85:398–409.

Geman S, Geman D. 1984. Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. IEEE Trans Pattern Analysis Machine Intell 6:721–741.

Gilks WR, Clayton DG, Spiegelhalter DJ, Best NG, McNeil AJ, Sharples LD, Kirby AJ. 1993. Modelling complexity: Applications of Gibbs sampling in medicine. J R Stat Soc 55:39–102.

Hastie TJ, Tibshirani RJ, eds. 1990. Generalized Additive Models. Chapman & Hall, New York NY.

Hastings WK. 1970. Monte Carlo sampling methods using Markov chains and their applications. Biometrika 57:97–109.

Kelsall JE, Samet JM, Zeger SL, Xu J. 1997. Air pollution and mortality in Philadelphia, 1974–1988. Am J Epidemiol 146:750–762.

Korrick SA, Neas LM, Dockery DW, Gold DR, Allen GA, Hill LB, Kimball KD, Rosner BA, Speizer FE. 1998. Effects of ozone and other pollutants on the pulmonary function of adult hikers. Environ Health Perspect 106:93–99.

Le Cam L, Yang GL, eds. 1990. Asymptotics in Statistics. Some Basic Concepts. Springer-Verlag, New York NY.

Liang KY, Zeger SL. 1986. Longitudinal data analysis using generalized linear models. Biometrika 73(1):13–22.

Lindley DV, Smith AFM. 1972. Bayes estimates for the linear model (with discussion). J R Stat Soc 34:1–41.

Meiring W, Guttorp P, Sampson P. 1997. Space-time estimation of grid-cell hourly ozone levels for assessment of a deterministic model. Technical Report. Statistics Department, University of Washington, Seattle WA.

Morris CN, Normand S-L. 1992. Hierarchical models for combining information and for meta-analysis. In: Bayesian Statistics 4. Proceedings of the Fourth Valencia International Meeting, Valencia, Spain. Oxford University Press, Oxford, England.

Neyman J, Scott EL. 1960. Correction for bias introduced by a transformation of variables. Ann Math Stat 31:643–655.

Normand STL, Glickman ME, Gatsonis C. 1997. Statistical methods for profiling providers of medical care: Issues and applications. J Am Stat Assoc 92:803–814.

Raftery A, Lewis S. 1992. How many iterations in the Gibbs sampler? In: Bayesian Statistics 4. Proceedings of the Fourth Valencia International Meeting (Bernardo JM, Berger JO, Dawid AP, Smith AFM, eds). Oxford University Press, Oxford, England.

Samet JM, Zeger SL, Berhane K. 1995. The association of mortality and particulate air pollution. In: Particulate Air Pollution and Daily Morality: Replication and Validation of Selected Studies, The Phase I.A Report of the Particle Epidemiology Evaluation Project. Health Effects Institute, Cambridge MA.

Samet J, Zeger S, Kelsall J, Xu J, Kalkstein L. 1997. Air pollution, weather, and mortality in Philadelphia 1973–1988. In: Particulate Air Pollution and Daily mortality: Analyses of the Effects of Weather and Multiple Air Pollutants, The Phase I.B Report of the Particle Epidemiology Evaluation Project. Health Effects Institute, Cambridge MA.

Schwartz J. 1993. Air pollution and daily mortality in Birmingham, Alabama. Am J Epidemiol 137:1136–1147.

Silliman NP. 1997. Hierarchical selection models with applications in meta-analysis. J Am Stat Assoc 92:926–936.

Skene AM, Wakefield JC. 1990. Hierarchical models for multicentre binary response studies. Stat Med 9:919–929.

Smith L, Davis J, Sacks J, Speckman P, Styer P. 1997. Assessing the human health risk of atmospheric particles. In: Statistics and the Environment. Proceedings from the 1997 Joint Statistical Meetings, Anaheim CA. American Statistical Association, Alexandria VA.

Smith R, Davis J, Sacks J, Speckman P, Styer P. 1998. Air Pollution and Mortality in Birmingham, Alabama: A Reappraisal.

Technical Report. National Institute of Statistical Science, Research Triangle Park NC.

Smith TC, Spiegelhalter DJ, Thomas A. 1995. Bayesian approaches to random-effects meta-analysis: A comparative study. Stat Med 14:2685–2699.

Tierney L. 1994. Markov chains for exploring posterior distributors. Ann Stat 22:1701–1762.

Wakefield J, Morris S. 1998. Spatial dependence and errors-in-variables in environmental epidemiology: Investigating the relationship between magnesium and myocardial infarction. In: Bayesian Statistics 6. Proceedings of the Sixth International Valencia Meeting, Valencia, Spain. Oxford University Press, Oxford, England.

Wickley C, Berliner M, Cressie N. 1997. Hierarchical Bayesian Space-Time Models. Technical Report. Iowa State University, Ames IA.

INTRODUCTION

The National Morbidity, Mortality, and Air Pollution Study (NMMAPS)* was designed to address uncertainties regarding the association between particulate air pollution and daily mortality. It was initiated in 1996, as the US Environmental Protection Agency (EPA) was reviewing the National Ambient Air Quality Standards (NAAQSs) for particulate matter (PM). The EPA PM Criteria Document (1996a) described studies from cities around the world that observed excess daily mortality associated with PM. Some observers, however, pointed to differences among the results from individual locations as suggesting that the association between PM and mortality is not one of cause and effect (Moolgavkar and Luebeck 1996). They observed that few studies had adequately addressed the possibility of confounding or effect modification by the non-PM pollutants, and when these issues were addressed, there was some evidence of an impact on the PM-mortality association (Moolgavkar et al 1995; Gamble and Lewis 1996). Thurston and Kinney (1995) observed that various statistical approaches to modeling of time-series data produced different results in the same data sets, and therefore the variation in methods might be an alternative explanation for differences in findings between cities.

Lipfert and Wyzga (1997) posited that errors in the measurement of air pollution exposure, attributable to both instrumental error and presumed lack of correlation between personal exposure and centrally measured ambient concentrations, made it difficult to attribute any observed excess in mortality to PM. McMichael and colleagues (1998) and Lipfert and Wyzga (1995) noted that even if the findings from current studies did indicate that excess daily mortality was caused by PM, the magnitude of reduction in life expectancy was unclear. If all the excess risk was confined to those frail persons who would have lived only for a few days longer if unexposed (a slight shortening of lifespan known as *mortality displacement* or *harvesting*), then the overall public health impact might not be as great. The EPA noted this uncertainty as well, but the results of two US prospective cohort studies (Dockery et al 1993; Pope et al 1995) were cited as evidence of larger reductions in life expectancy (EPA 1996b).

To conduct a nationwide multicity study, and to address the issues of measurement error and mortality displacement, new analytic methods were needed. NMMAPS Part I, *Methods and Methodologic Issues,* presents 5 sections that describe statistical methods that Dr Jonathan Samet and his colleagues developed to address exposure measurement error, mortality displacement, and model building for evaluating mortality in multiple cities and for multiple pollutants.† The methods for the multicity analysis are applied in NMMAPS Part II, *Morbidity, Mortality, and Air Pollution in the United States*, which presents results of analyses of overall and cause-specific mortality associated with PM and other pollutants in the 20 largest US cities and with PM alone in the 90 largest cities.

The original objectives of NMMAPS are well delineated in the Investigators' Report. NMMAPS Part I meets the objectives of methods development as described in the Overview. NMMAPS Part II addresses many of the remaining objectives, including an application of the multicity models developed in Part I, using data from the EPA, National Center for Health Statistics (NCHS), Health Care Financing Administration (HCFA), US Census, and National Weather Bureau. Also, the investigators are continuing to analyze the database, including analyses that combine morbidity and mortality data in several of the same cities considered in Part II; the findings will be published as a separate report.

Development of the methods in this report, which consider and examine issues of exposure measurement error and mortality displacement, has advanced the field of air pollution epidemiology in a major way, as has the application of the hierarchical models to analyze multiple cities and pollutants. Each of these issues is discussed below.

EXPOSURE MEASUREMENT ERROR

SCIENTIFIC BACKGROUND

NMMAPS Part I developed a conceptual framework and analytic approaches for considering the effect of exposure

* A list of abbreviations and other terms appears on page 13.

† Dr Jonathan Samet's investigation, *The National Morbidity, Mortality, and Air Pollution Study,* which will generate several reports, began in December 1996 and has cost about $700,000 to date. Part I of the Investigators' Report from Dr Samet and colleagues was received for review in May 1999. A revised report, received in September 1999, was accepted for publication in November 1999. During the review process, the HEI Review Committee and the investigators had the opportunity to exchange comments and to clarify issues in the Investigators' Report and in the Review Committee's Commentary.

measurement error on relative risk estimates. In epidemiologic studies, errors in measuring, quantifying, or classifying exposure can affect estimates of the association between exposure and health outcomes. Exposure measurement error not only includes any errors resulting from the measurement instrument, but also considers the error in assigning an individual's exposure based on instruments some distance away from each individual. Statisticians have investigated measurement error for at least half a century and have identified methods to overcome it. Statistical theory shows that in many situations the effect of exposure measurement error is to bias the estimate of association toward the null—that is, to lessen the estimated magnitude of any effect. Therefore, in these circumstances, the magnitude of the true unbiased effect would be larger than the observed (biased) effect estimate. This bias toward the null is not always the case, however.

One situation in which the bias is not towards the null occurs when a single measured value is taken to represent exposure in many individuals, whose average true exposure is equal to this single measured value. The resulting measurement error (called *Berkson error*) causes little or no bias in the effect estimates. Perhaps most worrisome are errors in measuring a variable that is confounding the association of interest, which can bias the effect estimate of that association in either direction. Error in measuring the variable of interest, if this error is correlated with an unmeasured or poorly measured confounding variable, can also lead to bias in either direction.

The impact of measurement error in ecologic time-series studies of air pollution has not been well studied. Thus, the generic criticism of these studies is that, because ambient air pollution is a poor measure of individuals' personal exposures, any positive results are likely to be spurious. Low correlations in studies comparing personal and ambient exposures appear to support such concerns. These concerns have been countered by findings that changes in pollutant concentrations at fixed central sites over time correlate well with changes in individually measured concentrations over time. Whereas ad hoc approaches have been used to address the effects of measurement error in these time-series studies, this report provides the first systematic consideration of how exposure measurement error might affect risk estimates between measures of PM less than 10 mg in aerodynamic diameter (PM_{10}) and mortality.

TECHNICAL EVALUATION

Section 1

In section 1, Zeger and colleagues establish a conceptual framework for considering exposure measurement error. The authors review the basic theory of measurement error effects, emphasizing the distinction between classical error, which causes bias in measures of association in most situations, and Berkson error, which causes little or no bias. This distinction provides the key to the framework proposed for the time-series context, where error is decomposed into 3 components, one of which is demonstrated to be of the Berkson type and therefore should not bias the association under investigation.

The 3 components of exposure measurement error discussed by the authors include differences between (1) individual exposure and average of individual exposures, (2) average individual exposure and true ambient exposure, and (3) measured ambient exposure and true ambient exposure. The components are primarily conceptual because the true exposures cannot easily be measured. The major Berkson error component is the difference between an individual's actual exposure to a particular pollutant and the average individual exposures of everyone in the city of interest. Strictly, this is a weighted average, but the authors convincingly show that the weights are usually unimportant. Therefore, if the true average population exposure were known and used in the time-series regression, the estimate of increment in risk per unit exposure would apply without bias to both individuals' exposures as well as the population average. The average is not known, however, and it is the ambient exposure measured by one or a few monitors that is generally used. The difference between the monitor measurements and the average personal exposure is the remaining error [components (2) and (3)] and is not of the Berkson type, and thus is likely to introduce bias in the effect estimate.

The authors further decompose this remaining error into the last 2 components: the difference between the average personal exposure and the true ambient exposure, and the difference between the true ambient exposure and the measured ambient exposure. It is not entirely clear what the true ambient exposure is, or how the further decomposition is useful. The authors argue, however, that the difference between the true and measured ambient exposure is likely to be of the Berkson type (introducing no bias). This would leave only the difference between average personal and true ambient exposure as likely to cause bias. The reasons for believing the third component (the difference between measured ambient and true ambient exposure) to be of the Berkson type are not clear, however. As the distinction is not used further, the issue appears to be academic; and this extra component may unnecessarily complicate the model.

The approach suggested by the investigators for correcting estimates of pollution-mortality relationships

broadly follows the relatively straightforward statistical method known as *regression calibration* (Carrol et al 1995), which assumes availability of data on both the error-prone daily ambient exposure measurements and personal exposure measurements for some persons on some days. These data are used to calibrate, that is, adjust, the ambient exposure measures by estimating from a regression model the change in average personal exposures corresponding to a unit change in ambient exposures. Once this calibration factor is known, the estimated change in mortality per unit change in ambient exposures can be corrected so that they apply to changes in personal exposures.

In section 1, the regression calibration is applied to data from the Particle Total Exposure Assessment Methodology (PTEAM) personal exposure study (Ozkaynak et al 1996) that included personal exposure measurements in Riverside, California, and ambient measures for the same days. The calibration regression found that a change of 1 $\mu g/m^3$ PM_{10} in the ambient measure was associated with a change of 0.60 $\mu g/m^3$ PM_{10} in average personal measures. Since the time-series analysis provides an estimate of effect for ambient exposures, the calibration factor for estimating the effect of personal exposure from an observed effect of ambient exposure, in this example, is 1/0.60 = 1.67. That is, the error introduced by using ambient exposure instead of personal exposure in estimating change in mortality per unit exposure can be adjusted, and, in this example, the adjustment results in a greater risk of mortality.

Using this method to estimate confidence intervals (CIs) for this risk estimate underestimates uncertainty, however, as the CIs do not reflect the imprecision in the calibration slope estimate (0.60). In section 1, simulations produce CIs without this limitation; the simulations in the example data were very similar to the calibration slope estimate. The simulation approach has the flexibility to extend to more complex situations, as illustrated in section 2. Simpler (but less easily extended) alternative methods for obtaining correct CIs are available (Rosner et al 1992; Carrol et al 1995), but were not considered in this report.

The regression calibration example in section 1 focuses on error in measuring exposure to a single pollutant. The authors, however, consider the theoretical impact of error in confounding variables and calculate bias under the classical error model with 2 exposure variables. These calculations confirmed conclusions from a similar exercise (Armstrong et al 1989) showing that, when 2 variables are correlated, errors in one can affect both regression coefficients, and that the effect on the regression coefficient of the other variable can be in either direction. When the 2 variables are measured with error and the errors are correlated, this cross contamination can be particularly pronounced. The authors adopted a simplified model for these calculations, and they reasonably argue that, even if this model is imperfect, the results give insight into the likely major consequences of exposure measurement errors.

Given the recognition that the relationship between observed ambient pollution and mortality is a biased estimate (usually underestimate) of the relationship between personal exposure and mortality, it becomes important to know which exposure-mortality relationship is useful for a specified purpose. Section 1 includes a discussion of this question. Because pollutant-level regulations are pertinent to ambient exposures, there is a prima facie argument that it is the (uncorrected) relationship between ambient exposure and mortality that is relevant for regulation. In this case, however, the distinction between measured and true ambient exposure, if clearly defined, could be important. The relationship between personal exposure and mortality remains important from a scientific and public health perspective, but it is less relevant to regulators. Specific interpretation of this relationship, however, raises questions discussed in the NMMAPS Part 1 report about the relative contribution to total exposure and toxicity of inhaled particles of those originating from ambient air and from indoor sources. Unfortunately, data that might be useful to answer these questions do not exist.

The development of the error framework, which appears sound, provides insight into the likely magnitude of effects of error, which overall is less than would be anticipated if a substantial component of the error was not of the Berkson type. This discussion also provides a statistical motivation for a method to correct estimates of pollution-mortality relationships for bias using data from small personal exposure studies with both personal and ambient exposure measurements.

An important consideration to drawing conclusions from this theoretical evaluation of effects of errors for actual ecologic time-series studies is the paucity of real-world data to examine the extent of errors and correlations between the variables and their errors (including cross-correlations). Complete reassurance that errors in measuring exposures could not have caused spurious associations (or more generally have led to bias toward the null) will have to await the collection of more exposure data. We agree with the conclusion of the authors that, for bias away from the null to be large, there would need to be high correlation between the true exposure and confounding variables, or between errors in these. We suggest that an additional theoretically possible cause of such bias is

substantial correlation between errors of measurement of exposure and other risk factors for death.

Section 2

Section 2 develops the framework established in section 1 by applying a form of the regression calibration bias-reduction method to situations where the information available on the relationship between the true and observed exposures has a hierarchical or grouped structure. The hierarchical structure is the result of data provided from studies conducted in several different cities. The data for each city included personal exposure measurements made on several individuals on several days, which are compared with the measured ambient exposure levels for those cities on the same days. The hierarchy comprises individual persons within cities and cities within the totality of the data set, allowing ambient-to-personal regression calibration slopes to be calculated for each city. To estimate the regression calibration slope for a new city, some sort of average of the city-specific slopes is needed. An appropriate statistical model should reflect uncertainty in the estimate of the calibration line for each city and also variation in calibration slopes between cities.

Section 2 uses a Bayesian modeling approach and the computational method called Markov chain Monte Carlo (MCMC). The Bayesian MCMC method is well established and has been used to address other regression problems with measurement error (Richardson and Gilks 1993), but its application to ecologic time-series studies with measurement error is new. An advantage of this approach is that it can incorporate complexities into the modeling that would cause problems in other methods. Thus, the log-linear Poisson regression model for time series could be linked to the hierarchical validity data without the need for simplifying assumptions that might be inappropriate.

Results from Bayesian analyses are in a different form than those from more familiar methods. For example, in the Bayesian analysis, the posterior mean and 95% posterior support intervals correspond to the more familiar point estimate of effect (such as relative risk) and 95% CI. This section also refers to the posterior interquartile range (IQR), which is roughly equivalent to a 50% CI.

As an example of the proposed method, time-series data from Baltimore are used to make a measurement error corrected estimate of change in mortality per 10 $\mu g/m^3$. Information on the relationship between ambient exposures and average personal exposure measurements was obtained from personal exposure studies in 5 cities. The average calibration slope (change in personal exposure per unit change in ambient exposure) across these 5 cities was estimated as 0.54 (compared with 0.60 found in Riverside, California, in section 1). However, individual city slopes varied among cities from 0.33 to 0.72, which is more than can be explained by chance. The posterior mean increment in mortality per 10 $\mu g/m^3$ ambient exposure was 0.9% (IQR, 0.67% to 1.12%). The corresponding figures for increment in mortality per 10 $\mu g/m^3$ personal exposure, using the information from the calibration slopes, were 1.44% (IQR, 0.94% to 1.88%).

For comparison, the simple form of regression calibration described in section 1 was also used. The results of the regression calibration were similar to those from the Bayesian MCMC hierarchical model, but since the simple estimate did not reflect uncertainty in the calibration slope, the 50% CI was somewhat narrower than the IQR. Some analyses of sensitivity of conclusions to model assumptions were made, which were broadly reassuring that specific assumptions were not critical.

The method described and illustrated in section 2 reflects the complexity of the context more fully than alternative methods, and it is flexible enough to extend relatively easily to incorporate details of data structures that may differ somewhat from those considered in the section. These strengths come at the cost of complexity in analysis and model assessment. Using the data on which the methods were illustrated, the simple and more complex methods gave similar results, but this may not always be so. In future considerations of exposure measurement error, methods of intermediate complexity may be worth considering in addition to the extremes of simplicity and sophistication described in this section.

CONCLUSIONS

Sections 1 and 2 advance our understanding of the effects of error in measuring air pollution in time-series studies. They present a theoretical model to test systematically what effect the relationship between personal exposures and ambient exposures might have on the observed increase in mortality associated with PM_{10}. These sections identify the likely direction of bias due to typical error (toward the null) and offer practical methods for its reduction. Lack of available information on the magnitude of measurement error means that the use of these methods to obtain estimates of the magnitude of bias (currently estimated to reduce the magnitude of the association by about half) is limited. They should be considered illustrative and tentatively suggestive of typical values rather than definitive. Studies relying on poor measures of ambient concentration will experience greater bias than those with good ambient measurements. Theoretical considerations of errors of measurement in multipollutant models and in confounding variables suggest that only highly correlated

variables and/or errors would produce substantial bias different from that described above. General absence of measured data on these errors and correlations, however, precludes confirmation that such conditions are not present. More data are being obtained in HEI, EPA, and other studies that should lead to more confident and specific conclusions.

MORTALITY DISPLACEMENT

SCIENTIFIC BACKGROUND

Mortality displacement, also referred to as *harvesting*, refers to premature death in the frailest members of a population—individuals who are near death die only slightly sooner than they would otherwise. If mortality displacement occurs, other, nonfatal, health consequences would be expected among less frail individuals. One might expect an acute increase in deaths among the frailest people in the population to signal a need for serious intervention to prevent further mortality and morbidity.

This idea of mortality displacement received empirical support from the observation that, for a period following the London Fog episode of very high air pollution levels in 1952, fewer deaths were observed than would have normally been expected. This was presumably due to the depletion during the pollution episode of a pool of susceptible individuals who would have died only a few days later if not for the London Fog. Recently, McMichael and colleagues (1998) and Lipfert and Wyzga (1995) have stated that the excess daily mortality measured in the time-series studies might be only, or largely, due to such short-term mortality displacement.

Drs L Cifuentes and LB Lave (Carnegie-Mellon University, Pittsburgh PA, unpublished data, 1996) demonstrated short-term mortality displacement by estimating expected deaths in Philadelphia across a 6-year period and then comparing this estimate to actual deaths that occurred during episodes of elevated total suspended particles (TSP) that lasted 3 days. The authors found that, at the beginning of the episode, the number of deaths was higher than expected when they compared it with the estimated average death rate. At the end of the episode, the authors found that the number of deaths was lower than expected. The excess deaths found were in addition to the excess deaths estimated by the Poisson regression model. The authors suggested the observed short-term excess of death at the beginning of the TSP episode was due to exposure to elevated TSP. Dr Claudia Spix (unpublished observations, 1998) has also developed modeling methods to consider how short-term mortality displacement might affect estimates of air pollution and mortality.

In this report, mortality displacement as described by Zeger and colleagues in section 3 involves 3 steps. First, the investigators proposed and investigated a 2-compartment model for mortality displacement, concluding that mortality displacement will only reflect itself on short-term time scales. Second, they evaluated a time-series model of the mortality–air pollution association that estimates the pollution relative risk at different time scales (year, season, month, day), referred to as the *frequency domain log-linear regression approach*. Taking the above 2 steps into account, the investigators propose a mortality displacement–resistant estimator that "sets aside" the short-term association that is subject to the influence of mortality displacement.

If short-term mortality displacement were the only reason for the association seen between mortality and pollution, effect estimates for long time periods (eg, a year) would show no effect. Therefore the authors conclude that, by examining the regressions using longer time periods, the pollution mortality association can be estimated without being biased by short-term mortality displacement. Schwartz (section 4) uses methods conceptually similar to those of Zeger and colleagues and applies them to studying total, chronic obstructive pulmonary disease (COPD), pneumonia, and ischemic heart disease (IHD) mortality in Boston.

TECHNICAL EVALUATION

In the analysis of the association between mortality and PM in section 3, the fact that mortality displacement may be occurring is not the main issue. Rather, the authors propose a method for estimating the air pollution mortality association that is insensitive to any mortality displacement effect.

The investigators first examine consequences of mortality displacement using simulation techniques similar to the ones used by previous investigators, assuming that individuals move from a healthy state to a frail (vulnerable) state, and are then susceptible to death. They demonstrate via computer simulations that mortality displacement occurs in the short term, by varying the average amount of time spent (between 3, 30, and 300 days) in the frail state, called the mean residence time (MRT). The simulations showed that if mortality displacement is the only source of the mortality–air pollution association, this association is near the null value at twice the MRT. This was true for each of the scenarios (3, 30, and 300 days).

The next step was to look at the association between air pollution and mortality using different time scales; that is, the data are expressed across year, season, month, week, and day. Regression techniques are applied to produce a regression coefficient that quantifies the association between mortality and pollution for a particular time scale.

If only mortality displacement accounts for the association between air pollution and mortality, then at very low frequencies (time scales varying over a long period, such as a year), no mortality displacement effect will result. At very high frequencies (time scales varying over short periods), such as a day, complete mortality displacement effect would be the result. Somewhere in between, a mortality displacement–resistant estimator can be determined. This estimator will then ignore all coefficients that correspond to periods shorter than twice the MRT in the frail group.

Data from Philadelphia are used in an example to assess the association between air pollution and mortality that is independent of mortality displacement. Figure 3, section 3, gives the estimates of mortality log relative risk associated with exposure to TSP over time. The pattern is different from that expected if mortality displacement is the sole explanation of this association. The figure shows that at low frequency (that is, long time periods, such as a year) there is an association between exposure to TSP and mortality that is different from 0 and is positive. At high frequencies (or shorter time periods, such as a day) there is a decrease in the association that comes close to 0. This is exactly the opposite of what is expected if mortality displacement is the only explanation of the association.

Section 4 similarly examines data from Boston over different time scales (or frequencies) for the association between air pollution levels and death rate at each of the different time scales. Here the author expected to find associations partly reflecting the effects of mortality displacement at the shorter time scales. Associations are also found at the longer time scales, however, which would suggest that death is being advanced by more than a few days.

The investigator separated the time-series data of daily deaths, air pollution, and weather into components of different wavelengths, or frequencies. These components are the low-frequency or long time-scale component, which will account for seasonal and annual trends in mortality; the moderate frequencies or mid-range time-scale component, which correspond to a time scale of a couple of months; and the high-frequency or short time-scale component (days), which is proposed to demonstrate mortality displacement.

Having separated out the mortality displacement and season effects, associations between air pollution and mortality that are not biased from these other factors can be investigated. The midscale range is defined at different time scales (frequency ranges) of 15, 30, 45, and 60 days. Associations between air pollution and mortality of different subgroups (COPD, pneumonia, and IHD) are investigated to establish time frames of impact—that is, where the largest impact on the association is seen. For COPD it seems to be a matter of days; for pneumonia and IHD it appears that exposures on the order of weeks or months have a significant impact.

One potential concern with interpreting the effects at the longer mortality displacement–resistant time scales is determining whether they are confounded by long-term time trends in the data. It is clear that the choice of frequencies removed from the data can have a substantial impact on estimates of effect (Cakmak et al 1998). If such confounding is present, then it is difficult to make conclusions about the implications of observed large effects at these longer time scales on the role of mortality displacement. In selecting the time scales of interest, one must somehow choose those that resist the effect of mortality displacement while still adequately controlling for effects of long-term temporal trends or cycles. We need to have assurance that longer-term time trends or cycles are adequately controlled before the findings of these mortality displacement sections can be taken as definitively supporting the intended interpretation.

Other issues that complicate interpretation of the observed estimates of effect at the larger time scales are those of cumulative effects, chronic effects, and lagged effects. Cumulative and chronic effects would require a certain pattern or period of exposure before they occur. Cumulative effects require that the condition worsens or is more likely to occur with each successive exposure. Chronic effects are similar to cumulative effects, but the term is more regularly used in association with an illness or disease that is simply long term regardless of exposure duration. A lagged effect in this context refers to an acute effect that follows the exposure by a day or more. The implications of the observed effects at the longer time scales for cumulative, chronic, and lagged effects, assuming these effects are not due to confounding, are not clear. Presumably there is little implication for lagged effects, since one can still address issues of lags using the longer time scales, as the investigators did. What is not clear is whether effects at the longer time scales imply something about cumulative or chronic effects. As more experience is gained with the approaches taken by the investigators, more insight into these issues may be gained as well.

The pollutant that might be associated with the short-term displacement is an additional consideration when

interpreting the findings. PM was the only component of the air pollution mix that was assessed in both sections on mortality displacement, and it has not been determined whether the short-term displacement reflects an effect of PM, an effect of another closely correlated air pollutant, or a more general effect of the air pollution mix.

Although sections 3 and 4 take conceptually similar approaches to addressing the issue of mortality displacement, there are some differences. The principal difference is estimating effects over a continuous time scale versus estimating effects at a few discrete time scales. One advantage to the continuous time scale is the ability to display estimates over the entire frequency spectrum. One disadvantage perhaps is that the procedure seems less transparent to those less expert in the statistical methods.

A second difference between the 2 sections is the focus on total mortality for Philadelphia (section 3), whereas cause-of-death categories were also evaluated for Boston (section 4). For Boston, effects were observed at the shorter time scales for COPD, but at the longer time scales for pneumonia and IHD. An attempt was made to interpret these findings, but questions regarding the interpretation remain. It could be argued that it would be more plausible for pneumonia and cardiovascular deaths, rather than deaths due to COPD, to exhibit the effects at short time scales. Here, as in the interpretation of the effect estimates at long time scales, however, it becomes easy to confuse cumulative, chronic, and lagged effects. Regardless, it will be interesting to see whether the patterns of effect at the various time scales by cause of death are consistent in Boston and Philadelphia, as has been proposed.

CONCLUSIONS

The investigators' examination of the role of mortality displacement using 2 different but related statistical methods is an original approach to evaluating whether more than a short-term displacement of mortality is occurring. The approach is convincing with a few qualifications. First, because PM was the only component of the air pollution mix that was assessed, it has not been determined whether this phenomenon is reflecting an effect of PM, an effect of another closely correlated air pollutant, or a more general effect of the air pollution mix. Second, given our lack of experience with these methods, it is also possible that use of time scales longer than a few days required by the techniques reintroduces bias from long-term time trends that influences the estimates of effects beyond a few days. Third, the approach does not allow a simultaneous assessment of the relative impact on mortality of both short-term displacement, if one is present, and longer-term effects.

As the authors have proposed, we need a more formal comparison of the methods described in the 2 sections, both to each other and to other approaches, to further our understanding of this issue. Given that this is the first application of this approach, replication of the findings in other settings is also needed; the general similarities in the findings from Philadelphia and Boston suggest that the findings could be reproduced in other settings.

NMMAPS METHODS

SCIENTIFIC BACKGROUND

Many components of ambient air pollution often derive from the same (combustion) sources, and therefore are at least moderately correlated over time in most urban areas. For this reason, epidemiologic studies of daily mortality in single cities are limited in their ability to estimate the independent effects of these individual components (Samet et al 1997). NMMAPS is based on the assumption that a multisite study that spans geographic areas differing with respect to the concentrations of the various components of ambient air pollution will provide better estimates of the independent and combined effects of each component than will combining several single-site studies after studies are completed. In addition, a multicity study offers the opportunity to explore characteristics of the environment or the population that might modify the effect of air pollution on daily mortality. Although these insights did not originate with NMMAPS, the investigators have applied them in new ways that are intended to improve on earlier approaches.

Consistency of effects over many different geographic areas can be useful in establishing evidence of an association, after proper adjustments have been made to account for differences in pollution profiles and population characteristics. One approach to examining results from several cities, combined in one analysis, has been the meta-analytic approach, which is a statistical method that provides a summary statistic of association using findings from available studies.

NMMAPS also combines results from individual cities. The NMMAPS approach differs from the previously used meta-analytic approach, however, in 2 potentially important ways. First, NMMAPS is based on a national air monitoring network, the Aerometric Information Retrieval System (AIRS) database, maintained by the EPA. NMMAPS sites were selected from this sampling frame in an explicit and prespecified fashion, as contrasted to the approach used in previous meta-analyses where sites were

not selected in a specified way. This sampling method alleviates concerns that selection or publication bias (the likelihood that studies with positive findings will be published more frequently than negative studies) will influence the reporting of NMMAPS findings in the scientific literature.

Second, in NMMAPS the multisite data are combined using a 2-stage Bayesian hierarchical modeling approach rather than the more informal meta-analytic methods applied previously. The first stage uses a standardized approach to estimate city-specific relative risks using log-linear models, employing adaptive smoothing techniques to adjust for time trends and potential confounders. The city-specific models build on extensive prior experience gained in the Particle Epidemiology Evaluation Project (PEEP) study concerning how best to control for confounding factors such as weather, day of the week, seasonality, and long-term trends in mortality (Samet et al 1997).

Semiparametric smoothing techniques are used separately by city to adjust for fluctuations in possible effects on mortality from weather, day of the week, age-specific long-term trends, and other factors in a standardized fashion. The advantage of these semiparametric methods of adjustment is that they avoid the need to make modeling assumptions such as linearity, and they are therefore a good choice to use when adjusting for factors such as seasonality and weather, where the exact shape of the exposure-response relationship is unknown. The application of these methods allows the data from each city to be used in a flexible and adaptive way best suited to that particular location.

The second stage of the analysis combines the pollution-mortality regression coefficients across the cities using a Bayesian hierarchical modeling approach. This model provides an overall combined estimate of the effect and a unified method for examining whether other city-specific factors, including sociodemographic and meteorologic variables, may help explain heterogeneity of pollutant effects on mortality across cities. In epidemiologic terms, variables found to explain variability in the city-specific regression coefficients are effect modifiers. This hierarchical approach also takes into account the heterogeneity of effects across cities in computing the overall summary measure of effect.

TECHNICAL EVALUATION

Section 5 develops the statistical model and methodologic framework to be used to model the relationship between pollutants and mortality and to combine evidence across multiple cities. The authors use a 2-stage Bayesian hierarchical modeling approach that provides a flexible and comprehensive modeling strategy well suited to combining effects from multiple cities. Flexibility is built in to the Poisson log-linear modeling approach used for city-specific analyses in the first stage through the use of semiparametric generalized additive modeling that incorporates adaptive smoothing techniques. This smoothing approach is adaptive in the sense that it allows the relationship between mortality and key covariates that fluctuate over time, such as daily temperature, to vary in complex ways from city to city. The extent of smoothing is controlled by the degrees of freedom (*df*) of the smoother, and sensitivity analyses are used to explore the effect of increasing or decreasing the amount of smoothing allowed.

The goal of the first stage of the modeling is to provide an automated approach that allows a sufficiently flexible model to be fit to data from each individual city. The second-stage Bayesian hierarchical model provides a natural approach to combining results across cities to provide an overall estimate of effect. It also provides a method to explore factors (effect modifiers) that may in part explain city-to-city heterogeneity, that is, dissimilar associations between air pollution and mortality by city. Finally, the Bayesian approach also provides posterior estimates of effect for individual cities that borrow strength from the entire set of cities. That is, the posterior mean of the estimated effect for each city is a shrinkage estimate that takes into account the precision with which the regression coefficient is estimated for that city, and shrinks the estimate toward the overall combined mean. The amount of shrinkage is less for cities with precisely estimated regression coefficients and more for cities with poorly estimated coefficients. This approach may improve estimates of effects of pollutants for individual cities, particularly where the effects of pollution on mortality are poorly estimated because of small numbers of deaths. Proposed models as applied in this section do not incorporate substantial prior opinion about parameters of interest, so that results are not influenced to any substantial degree by possibly subjective prior information.

The general approach for modeling mortality over time within each city is well thought out. The model attempts to control for long-term mortality trends as well as short-term trends due, for example, to phenomena such as influenza epidemics. Temperature variables (temperature and dew point on the current day and average for the previous 3 days) and day of the week are also controlled for in the model. The model stratifies on 3 age groups (<65 years, 65–75, >75), and includes interactions of age with some other terms in the model (eg, short-term mortality trends are allowed to be age-specific). As mentioned, smoothing

techniques were used to model the long-term and short-term fluctuations in mortality, as well as the effects of temperature and dew point, making the models semiparametric in that sense.

Three separate analyses were run for each pollutant, examining effects of pollutant levels on the current day (lag 0), prior day (lag 1), and 2 days prior to death (lag 2). Unfortunately, because PM was measured only on every sixth day in some cities, it was not possible to fit models incorporating the effects of more than one lag (referred to as *distributed lag models*). The models were fit assuming independence of errors over time. Examination of the autocorrelation function of the standardized residuals indicated that correlations were very close to 0, suggesting that this was a reasonable assumption, as well as providing some reassurance that seasonality and short-term time trends were adequately controlled for in the model.

The second-stage model describes between-city variation of the true underlying city-specific effects in terms of selected city-specific covariates. It also provides a combined estimate of the overall effect of each pollutant (univariate model) or pair of pollutants (bivariate model). Several variants of the second-stage model are described, including a baseline model that assumes that the true city-specific rates are uncorrelated and 2 versions of a spatial model, which allow the true city-specific rates to be correlated geographically using 2 different distance measures in the correlation function. These three 2-stage models in combination provide sufficiently broad flexibility in the approach used to examine and describe city-to-city heterogeneity.

The sensitivity analyses (section 5, Table 4) are reassuring in that they indicate robustness of the findings to the choice of lag exposure time, prior distributions of parameters in the Bayesian framework, and choice of second-stage model (baseline univariate, baseline bivariate, or spatial). Sensitivity analyses were done (results not shown) where the *df* (smoothness) of the nonparametric functions modeling time, weather, and seasonality were varied over a fourfold range. Assuming that the default *df* are in the generally optimal region, the finding that the degree of smoothing had little effect on the regression coefficient of pollutants of primary interest was reassuring.

Fitting sufficiently general models that adequately control confounders to city-specific data and finding a consistent, adjusted effect across cities while exploring the robustness of results and fit of the model using sensitivity analyses and other methods are important for elucidating the associations. One potential issue with the Bayesian hierarchical approach is its perceived complexity among those not familiar with this relatively new application of this methodology. In this sense, comparison of the results obtained from this approach with simpler approaches such as random effects meta-analysis (DerSimonian and Laird 1986) would be of interest. Another potential limitation of the Bayesian hierarchical approach is the possible dependence of reported results on the prior distribution. The NMMAPS analysis, however, is based on no-information or low-information prior distributions, and their influence should be minimal.

CONCLUSIONS

The Bayesian hierarchical approach taken to multicity modeling is a flexible and comprehensive modeling approach for estimating effects of air pollutants in multiple cities, combining their effects across cities, and exploring factors leading to heterogeneity among cities.

Section 5 lays out a general methodology for carrying out multicity analyses. The example provided in the section illustrates the methodology, but is not intended to be a comprehensive analysis of the effects of air pollution and mortality. NMMAPS Part II will further explore factors that might explain heterogeneity of the pollutant-mortality association across cities. Section 5 also lays a foundation for addressing other important issues including (1) possible nonlinearity of the exposure-mortality relationship, (2) variation in the exposure relationships by season of the year (pollutant by season interactions), and (3) possible interactions of age with other exposure variables such as temperature or dew point.

REFERENCES

Armstrong BG, Whittemore AS, Howe GR. 1989. Analysis of case-control data with covariate measurement error: Application to diet and colon cancer. Stat Med 8:1151–1163.

Cakmak S, Burnett R, Krewski D. 1998. Adjusting for temporal variation in the analysis of parallel time series of health and environmental variables. J Expo Anal Environ Epidemiol 8:129–144.

Carroll RJ, Ruppert D, Stefanski LA. 1995. Measurement error in nonlinear models. Chapman & Hall, New York NY.

DerSimonian R, Laird N. 1986. Meta analysis in clinical trials. Control Clin Trials 7:177–188.

Dockery DW, Pope CA III, Xu X, Spengler JD, Ware JH, Fay ME, Ferris BG Jr, Speizer FE. 1993. An association between air pollution and mortality in six US cities. N Engl J Med 329:1753–1759.

Gamble JF, Lewis RJ. 1996. Health and respirable particulate (PM_{10}) air pollution: A causal or statistical association? Environ Health Perspect 104:838–850.

Lipfert FW, Wyzga RE. 1995. Air pollution and mortality: Issues and uncertainties. J Air Waste Manage Assoc 45: 949–966.

Lipfert FW, Wyzga RE. 1997. Air pollution and mortality: The implications of uncertainties in regression modeling and exposure measurement. J Air Waste Manage Assoc 47:517–523.

McMichael AJ, Anderson HR, Brunekreef B, Cohen AJ. 1998. Inappropriate use of daily mortality analyses to estimate longer-term effects of air pollution. Int J Epidemiol 27:450–453.

Moolgavkar SH, Luebeck EG. 1996. A critical review of the evidence on particulate air pollution and mortality. Epidemiology 7:420–428.

Moolgavkar SH, Luebeck EG, Hall TA, Anderson EL. 1995. Air pollution and daily mortality in Philadelphia. Epidemiology 6:476–484.

Ozkaynak H, Xue J, Spengler J, Wallace L, Pellizzari E, Jenkins P. 1996. Personal exposure to airborne particles and metals: Results from the Particle TEAM study in Riverside, California. J Expo Anal Environ Epidemiol 6:57–78.

Pope CA III, Thun MJ, Namboodiri MM, Dockery DW, Evans JS, Speizer FE, Heath CW Jr. 1995. Particulate air pollution as a predictor of mortality in a prospective study of US adults. Am J Respir Crit Care Med 151:669–674.

Richardson S, Gilks WR. 1993. A Bayesian approach to measurement error problems in epidemiology using conditional independence models. Am J Epidemiol 138:430–442.

Rosner B, Spiegelman D, Willett WC. 1992. Correction of logistic regression relative risk estimates and confidence intervals for random within-person measurement error. Am J Epidemiol 136:1400–1413.

Samet JM, Zeger DL, Kelsall JE, Xu J. 1997. Air pollution, weather, and mortality in Philadelphia 1973–1988. In: Particulate Air Pollution and Daily Mortality: Analyses of the Effects of Weather and Multiple Air Pollutants, The Phase I.B Report of the Particle Epidemiology Evaluation Project. Health Effects Institute, Cambridge MA.

Schwartz J. 1994. Air pollution and daily mortality: A review and meta-analysis. Environ Res 64:36–52.

Thurston DG, Kinney PL. 1995. Air pollution epidemiology: Considerations in time-series modeling. Inhalation Toxicol 7:71–83.

US Environmental Protection Agency. 1996a. Air Quality Criteria for Particulate Matter. EPA/600/P-95/001cF. Office of Research and Development, Washington DC.

US Environmental Protection Agency. 1996b. Review of the National Air Quality Standards for Particulate Matter: Policy Assessment of Scientific and Technical Information. OAQPS Staff Paper. EPA-452\R-96-013. Office of Air Quality Planning and Standards, Research Triangle Park NC.

RELATED HEI PUBLICATIONS: PARTICULATE MATTER

Report No*	Title	Principal Investigator	Publication Date
Research Reports			
40	Retention Modeling of Diesel Exhaust Particles in Rats and Humans	CP Yu	1991
45	The Effects of Exercise on Dose and Dose Distribution of Inhaled Automotive Pollutants	MT Kleinman	1991
68	Pulmonary Toxicity of Inhaled Diesel Exhaust and Carbon Black in Chronically Exposed Rats		
	Part I. Neoplastic and Nonneoplastic Lung Lesions	JL Mauderly	1995
	Part II. DNA Damage	K Randerath	1995
	Part III. Examination of Possible Target Genes	SA Belinsky	1995
75	Ozone Exposure and Daily Mortality in Mexico City: A Time-Series Analysis	DP Loomis	1996
83	Daily Changes in Oxygen Saturation and Pulse Rate Associated with Particulate Air Pollution and Barometric Pressure	DW Dockery	1999
86	Statistical Methods for Epidemiologic Studies of the Health Effects of Air Pollution	W Navidi	1999
91	Mechanisms of Morbidity and Mortality from Exposure to Ambient Air Particles	JJ Godleski	2000
93	Effects of Concentrated Ambient Particles in Rats and Hamsters: An Exploratory Study	T Gordon	2000
94	The National Morbidity, Mortality, and Air Pollution Study *Part II.* Morbidity and Mortality from Air Pollution in the United States	JM Samet	2000
95	Association of Particulate Matter Components with Daily Mortality and Morbidity in Urban Populations	M Lippmann	2000
96	Acute Pulmonary Effects of Ultrafine Particles in Rats and Mice	G Oberdörster	2000
…	A Case-Crossover Analysis of Fine Particulate Matter Air Pollution and Out-of-Hospital Sudden Cardiac Arrest (in press)	H Checkoway	2000
…	Identifying Subgroups of the General Population that May Be Susceptible to Short-Term Increases in Particulate Air Pollution: A Time-Series Study in Montreal, Quebec (in press)	MS Goldberg	2000
…	Daily Mortality and Fine and Ultrafine Particles in Erfurt, Germany *Part I.* Role of Particle Number and Particle Mass (in press)	H-E Wichmann	2000
Special Reports			
	Particulate Air Pollution and Daily Mortality: The Phase I Report of the Particle Epidemiology Evaluation Project		
	Phase I.A. Replication and Validation of Selected Studies		1995
	Phase I.B. Analyses of the Effects of Weather and Multiple Air Pollutants		1997
	Reanalysis of the Harvard Six Cities Study and the American Cancer Society Study of Particulate Air Pollution and Mortality		2000

continued

** Reports published since 1990. Copies of these reports can be obtained by contacting the Health Effects Institute, 955 Massachusetts Avenue, Cambridge MA 02139. Phone (617) 876-6700 FAX (617) 876-6709 E-mail pubs@healtheffects.org www.healtheffects.org*

RELATED HEI PUBLICATIONS

Research Reports

Report No.	Title	Principal Investigator	Publication Date
HEI Communications			
5	Formation and Characterization of Particles: Report of the 1996 HEI Workshop		1997
8	The Health Effects of Fine Particles: Key Questions and the 2003 Review		1999
HEI Program Summaries			
	Research on Particulate Matter		1999

Copies of these reports can be obtained by contacting the Health Effects Institute, 955 Massachusetts Avenue, Cambridge, MA 02139. Phone (617) 876-6700 FAX (617) 876-6709 E-mail pubs@healtheffects.org www.healtheffects.org

HEALTH
EFFECTS
INSTITUTE

955 Massachusetts Avenue
Cambridge MA 02139 USA
+1-617-876-6700
www.healtheffects.org

RESEARCH
REPORT

Number 94, Part I
May 2000